

PCT

WORLD INTELLECTUAL PROPERTY ORGANIZATION  
International Bureau



INTERNATIONAL APPLICATION PUBLISHED UNDER THE PATENT COOPERATION TREATY (PCT)

(51) International Patent Classification 5 : <b>C07K 13/00, C12N 15/12, 5/16 C12Q 1/00</b>	A1	(11) International Publication Number: <b>WO 93/04083</b> (43) International Publication Date: <b>4 March 1993 (04.03.93)</b>
(21) International Application Number: <b>PCT/US92/06903</b>		(74) Agents: WATT, Phillip, H. et al.; Fitch, Even, Tabin & Flannery, Room 900, 135 South LaSalle Street, Chicago IL 60603 (US).
(22) International Filing Date: <b>14 August 1992 (14.08.92)</b>		(81) Designated States: AT, AU, BB, BG, BR, CA, CH, CS, DE, DK, ES, FI, GB, HU, JP, KP, KR, LK, LU, MG, MN, MW, NL, NO, PL, RO, RU, SD, SE, US, European patent (AT, BE, CH, DE, DK, ES, FR, GB, GR, IE, IT, LU, MC, NL, SE), OAPI patent (BF, BJ, CF, CG, CI, CM, GA, GN, ML, MR, SN, TD, TG).
(30) Priority data: <b>745,206 868,354</b> 15 August 1991 (15.08.91) US 10 April 1992 (10.04.92) US		Published <i>With international search report.</i>
(60) Parent Application or Grant (63) Related by Continuation US Filed on <b>868,354 (CIP) 10 April 1992 (10.04.92)</b>		
(71) Applicant ( <i>for all designated States except US</i> ): THE SALK INSTITUTE BIOTECHNOLOGY/INDUSTRIAL ASSOCIATES, INC. [US/US]; 505 Coast Boulevard South, Suite 300, La Jolla, CA 92037 (US).		
(72) Inventors; and (75) Inventors/Applicants ( <i>for US only</i> ) : HARPOLD, Michael, Miller [US/US]; 1341 29th Street, San Diego, CA 92102 (US). ELLIS, Steven, Bradley [US/US]; 8939 Oviedo Street, San Diego, CA 92129 (US). WILLIAMS, Mark, Edward [US/US]; 6919 Pear Tree Drive, Carlsbad, CA 92009 (US). FELDMAN, Daniel, Howard [US/US]; 4621 Lamont, #9A, San Diego, CA 92109 (US). McCUE, Ann, Francis [US/US]; 6939 Colorado Avenue, La Mesa, CA 91942 (US). BRENNER, Robert [US/US]; 2507 Kenmore Court, Austin, TX 78703 (US).		

(54) Title: HUMAN CALCIUM CHANNEL COMPOSITIONS AND METHODS

(57) Abstract

Isolated DNA encoding each of human calcium channel  $\alpha_1$ -,  $\alpha_2$ -,  $\beta$ - and  $\gamma$ -subunits, including subunits that arise as splice variants of primary transcripts, is provided. Cells and vectors containing the DNA and methods for identifying compounds that modulate the activity of human calcium channels are also provided.

A0000179/2-66-MG  
10/088,257

**FOR THE PURPOSES OF INFORMATION ONLY**

**Codes used to identify States party to the PCT on the front pages of pamphlets publishing international applications under the PCT.**

AT	Austria	FI	Finland	MN	Mongolia
AU	Australia	FR	France	MR	Mauritania
BB	Barbados	GA	Gabon	MW	Malawi
BE	Belgium	GB	United Kingdom	NL	Netherlands
BF	Burkina Faso	CN	Guinea	NO	Norway
BC	Bulgaria	GR	Greece	NZ	New Zealand
BJ	Benin	HU	Hungary	PL	Poland
BR	Brazil	IE	Ireland	PT	Portugal
CA	Canada	IT	Italy	RO	Romania
CF	Central African Republic	JP	Japan	RU	Russian Federation
CG	Congo	KP	Democratic People's Republic of Korea	SD	Sudan
CH	Switzerland	KR	Republic of Korea	SE	Sweden
CI	Côte d'Ivoire	LI	Liechtenstein	SK	Slovak Republic
CM	Cameroon	LK	Sri Lanka	SN	Senegal
CS	Czechoslovakia	LU	Luxembourg	SU	Soviet Union
CZ	Czech Republic	MC	Monaco	TD	Chad
DE	Germany	MG	Madagascar	TG	Togo
DK	Denmark	MI	Mali	UA	Ukraine
ES	Spain			US	United States of America

- 1 -

### HUMAN CALCIUM CHANNEL COMPOSITIONS AND METHODS

This application is a continuation-in-part of United States Serial No. 07/868,354, filed April 10, 1992, which is a continuation-in-part of United States Serial No.

5 07/745,206, filed August 15, 1991, which is a continuation-in-part of United States Serial No. 07/620,250, filed November 30, 1990, which is a continuation-in-part of United States Serial No. 07/176,899, filed April 4, 1988, now abandoned, and is also a continuation-in-part of United  
10 States Serial No. 07/482,384, filed February 20, 1990, and is also a continuation-in-part of United States Serial No. 07/941,231, filed July 13, 1992, which in turn is a continuation of United States Serial No. 07/603,751, filed April 4, 1989, now abandoned.

15 **TECHNICAL FIELD**

The present invention relates to molecular biology and pharmacology. More particularly, the invention relates to calcium channel compositions and methods of making and using the same.

20 **BACKGROUND OF THE INVENTION**

Calcium channels are membrane-spanning, multi-subunit proteins that allow controlled entry of  $\text{Ca}^{2+}$  ions into cells from the extracellular fluid. Cells throughout the animal kingdom, and at least some bacterial, fungal and plant  
25 cells, possess one or more types of calcium channel.

The most common type of calcium channel is voltage dependent. "Opening" of a voltage-dependent channel to allow an influx of  $\text{Ca}^{2+}$  ions into the cells requires a depolarization to a certain level of the potential  
30 difference between the inside of the cell bearing the channel and the extracellular medium bathing the cell. The rate of influx of  $\text{Ca}^{2+}$  into the cell depends on this potential difference. All "excitable" cells in animals, such as neurons of the central nervous system (CNS),  
35 peripheral nerve cells and muscle cells, including those of skeletal muscles, cardiac muscles, and venous and arterial smooth muscles, have voltage-dependent calcium channels.

-2-

Multiple types of calcium channels have been identified in mammalian cells from various tissues, including skeletal muscle, cardiac muscle, lung, smooth muscle and brain, [see, e.g., Bean, B.P.(1989) *Ann. R v. Physiol.* 51:367-384 and Hess, P. (1990) *Ann. Rev. Neurosci.* 56:337]. The different types of calcium channels have been broadly categorized into four classes, L-, T-, N-, and P-type, distinguished by current kinetics, holding potential sensitivity and sensitivity to calcium channel agonists and antagonists.

Calcium channels are multisubunit proteins. For example, rabbit skeletal muscle calcium channel contains two large subunits, designated  $\alpha_1$  and  $\alpha_2$ , which have molecular weights between about 130 and about 200 kilodaltons ("kD"), and one to three different smaller subunits of less than about 60 kD in molecular weight. At least one of the larger subunits and possibly some of the smaller subunits are glycosylated. Some of the subunits are capable of being phosphorylated. The  $\alpha_1$  subunit has a molecular weight of about 150 to about 170 kD when analyzed by sodium dodecylsulfate (SDS)-polyacrylamide gel electrophoresis (PAGE) after isolation from mammalian muscle tissue and has specific binding sites for various 1,4-dihydropyridines (DHPs) and phenylalkylamines. Under non-reducing conditions (in the presence of N-ethylmaleimide), the  $\alpha_2$  subunit migrates in SDS-PAGE as a band corresponding to a molecular weight of about 160-190 kD. Upon reduction, a large fragment and smaller fragments are released. The  $\beta$  subunit of the rabbit skeletal muscle calcium channel is a phosphorylated protein that has a molecular weight of 52-65 kD as determined by SDS-PAGE analysis. This subunit is insensitive to reducing conditions. The  $\gamma$  subunit of the calcium channel, which is not observed in all purified preparations, appears to be a

-3-

glycoprotein with an apparent molecular weight of 30-33 kD, as determined by SDS-PAGE analysis.

In order to study calcium channel structure and function, large amounts of pure channel protein are needed.

5 Because of the complex nature of these multisubunit proteins, the varying concentrations of calcium channels in tissue sources of the protein, the presence of mixed populations of calcium channels in tissues, difficulties in obtaining tissues of interest, and the modifications of the  
10 native protein that can occur during the isolation procedure, it is extremely difficult to obtain large amounts of highly purified, completely intact calcium channel protein.

Characterization of a particular type of calcium  
15 channel by analysis of whole cells is severely restricted by the presence of mixed populations of different types of calcium channels in the majority of cells. Single-channel recording methods that are used to examine individual calcium channels do not reveal any information regarding  
20 the molecular structure or biochemical composition of the channel. Furthermore, in performing this type of analysis, the channel is isolated from other cellular constituents that might be important for natural functions and pharmacological interactions.

25 Characterization of the gene or genes encoding calcium channels provides another means of characterization of different types of calcium channels. The amino acid sequence determined from a complete nucleotide sequence of the coding region of a gene encoding a calcium channel  
30 protein represents the primary structure of the protein. Furthermore, secondary structure of the calcium channel protein and the relationship of the protein to the membrane may be predicted based on analysis of the primary structure. For instance, hydropathy plots of the  $\alpha_1$  subunit  
35 protein of the rabbit skeletal muscle calcium channel

-4-

indicate that it contains four internal repeats, each containing six putative transmembrane regions [Tanabe, T. et al. (1987) *Nature* 328:313].

The cDNA and corresponding amino acid sequences of the  
5  $\alpha_1$ ,  $\alpha_2$ ,  $\beta$  and  $\gamma$  subunits of the rabbit skeletal muscle  
calcium channel [see, Tanabe et al. (1987) *Nature* 328:313-  
318; International Application No. WO 89/09834, which is  
U.S. Application Serial No. 07/603,751, which is a  
continuation-in-part of U.S. Application Serial No.  
10 07/176,899; Ruth et al. (1989) *Science* 245:1115-1118; and  
U.S. Patent Application Serial No. 482,384, filed February  
20, 1990] have been determined. The cDNA and corresponding  
amino acid sequences of  $\alpha_1$  subunits of rabbit cardiac muscle  
[Mikami, A. et al. (1989) *Nature* 340:230-233] and lung  
15 [Biel, M. (1990) *FEBS Letters* 269:409-412] calcium channels  
have been determined.

In addition, a cDNA clone encoding a rabbit brain  
calcium channel (designated the BI channel) has been  
isolated [Mori, Y. et al. (1991) *Nature* 350:398-402].  
20 Partial cDNA clones encoding portions of several different  
subtypes, referred to as rat brain class A, B, C and D, of  
the calcium channel  $\alpha_1$  subunit have been isolated from rat  
brain cDNA libraries [Snutch, T. et al. (1990) *Proc. Natl.  
Acad. Sci. USA* 87:3391-3395]. More recently full-length  
25 rat brain class A [Starr, T. et al. (1991) *Proc. Natl.  
Acad. Sci. USA* 88:5621-5625] and class C [Snutch, T. et al.  
(1991) *Neuron* 7:45-57] cDNA clones have been isolated.  
Although the amino acid sequence encoded by the rat brain  
class C DNA is approximately 95% identical to that encoded  
30 by the rabbit cardiac muscle calcium channel  $\alpha_1$  subunit-  
encoding DNA, the amino acid sequence encoded by the rat  
brain class A DNA shares only 33% sequence identity with  
the amino acid sequence encoded by the rabbit skeletal or  
cardiac muscle  $\alpha_1$  subunit-encoding DNA. A cDNA clone  
35 encoding another rat brain calcium channel  $\alpha_1$  subunit has

-5-

also been obtained [Hui, A. et al. (1991) *Neuron* 7:35-44]. The amino acid sequence encoded by this clone is ~70% homologous to the proteins encoded by the rabbit skeletal and cardiac muscle calcium channel DNA. A cDNA clone 5 closely related to the rat brain class C  $\alpha_1$  subunit-encoding cDNA and sequences of partial cDNA clones closely related to other partial cDNA clones encoding apparently different calcium channel  $\alpha_1$  subunits have also been isolated [see Snutch, T. et al. (1991) *Neuron* 7:45-57; Perez-Reyes, E. et 10 al. (1990) *J. Biol. Chem.* 265:20430; and Hui, A. et al. (1991) *Neuron* 7:35-44]. DNA clones encoding other calcium channels have also been identified and isolated.

Expression of cDNA encoding calcium channel subunits has been achieved with several of the different rabbit or 15 rat  $\alpha_1$  subunit cDNA clones discussed above. Voltage-dependent calcium currents have been detected in murine L cells transfected with DNA encoding the rabbit skeletal muscle calcium channel  $\alpha_1$  subunit [Perez-Reyes et al. (1989) *Nature* 340:233-236 (1989)]. These currents were enhanced 20 in the presence of the calcium channel agonist Bay K 8644. Bay K 8644-sensitive Ba<sup>2+</sup> currents have been detected in oocytes injected with *in vitro* transcripts of the rabbit cardiac muscle calcium channel  $\alpha_1$  subunit cDNA [Mikami, A. et al. (1989) *Nature* 340:230-233]. These currents were 25 substantially reduced in the presence of the calcium channel antagonist nifedipine. Barium currents of an oocyte co-injected with RNA encoding the rabbit cardiac muscle calcium channel  $\alpha_1$  subunit and the RNA encoding the rabbit skeletal muscle calcium channel  $\alpha_2$  subunit were more 30 than 2-fold larger than those of oocytes injected with transcripts of the rabbit cardiac calcium channel  $\alpha_1$  subunit-encoding cDNA. Similar results were obtained when oocytes were co-injected with RNA encoding the rabbit lung calcium channel  $\alpha_1$  subunit and the rabbit skeletal muscle 35 calcium channel  $\alpha_2$  subunit. The barium current was greater

-6-

than that detected in oocytes injected only with RNA encoding the rabbit lung calcium channel  $\alpha_1$  subunit [Biel, M. et al. (1990) *FEBS Letters* 269:409-412]. Inward barium currents have been detected in oocytes injected with *in vitro* RNA transcripts encoding the rabbit brain BI channel [Mori et al. (1991) *Nature* 350:398-402]. These currents were increased by two orders of magnitude when *in vitro* transcripts of the rabbit skeletal muscle calcium channel  $\alpha_2$ ,  $\beta$ , or  $\alpha_2$ ,  $\beta$  and  $\gamma$  subunits were co-injected with transcripts of the BI-encoding cDNA. Barium currents in oocytes co-injected with transcripts encoding the BI channel and the rabbit skeletal muscle calcium channel  $\alpha_2$  and  $\beta$  were unaffected by the calcium channel antagonists nifedipine or  $\omega$ -CgTx and inhibited by Bay K 8644 and crud venom from *Agelenopsis aperta*.

The results of studies of recombinant expression of rabbit calcium channel  $\alpha_1$  subunit-encoding cDNA clones and transcripts of the cDNA clones indicate that the  $\alpha_1$  subunit forms the pore through which calcium enters cells. The relevance of the barium currents generated in these recombinant cells to the actual current generated by calcium channels containing as one component the respective  $\alpha_1$  subunits *in vivo* is unclear. In order to completely and accurately characterize and evaluate different calcium channel types, however, it is essential to examine the functional properties of recombinant channels containing all of the subunits as found *in vivo*. Although there has been limited success in expressing DNA encoding rabbit and rat calcium channel subunits, far less has been achieved with respect to human calcium channels. Little is known about human calcium channel structure and function and gene expression. An understanding of the structure and function of human calcium channels would permit identification of substances that, in some manner, modulate the activity of

-7-

calcium channels and that have potential for use in treating such disorders.

Because calcium channels are present in various tissues and have a central role in regulating intracellular calcium ion concentrations, they are implicated in a number of vital processes in animals, including neurotransmitter release, muscle contraction, pacemaker activity, and secretion of hormones and other substances. These processes appear to be involved in numerous human disorders, such as CNS and cardiovascular diseases. Calcium channels, thus, are also implicated in numerous disorders. A number of compounds useful for treating various cardiovascular diseases in animals, including humans, are thought to exert their beneficial effects by modulating functions of voltage-dependent calcium channels present in cardiac and/or vascular smooth muscle. Many of these compounds bind to calcium channels and block, or reduce the rate of, influx of  $\text{Ca}^{2+}$  into the cells in response to depolarization of the cell membrane.

An understanding of the pharmacology of compounds that interact with calcium channels in other organ systems, such as the CNS, may aid in the rational design of compounds that specifically interact with subtypes of human calcium channels to have desired therapeutic effects, such as in the treatment of neurodegenerative and cardiovascular disorders. Such understanding and the ability to rationally design therapeutically effective compounds, however, have been hampered by an inability to independently determine the types of human calcium channels and the molecular nature of individual subtypes, particularly in the CNS, and by the unavailability of pure preparations of specific channel subtypes to use for evaluation of the specificity of calcium channel-effecting compounds. Thus, identification of DNA encoding human calcium channel subunits and the use of such DNA for

-8-

expression of calcium channel subunits and functional calcium channels would aid in screening and designing therapeutically effective compounds.

Therefore, it is an object herein, to provide DNA 5 encoding specific calcium channel subunits and to provide eukaryotic cells bearing recombinant tissue-specific or subtype-specific calcium channels. It is also an object to provide assays for identification of potentially therapeutic compounds that act as calcium channel 10 antagonists and agonists.

#### SUMMARY OF THE INVENTION

Eukaryotic cells containing heterologous DNA encoding one or more calcium channel subunits, particularly human calcium channel subunits, or containing RNA transcripts of 15 DNA clones encoding one or more of the subunits are provided. In preferred embodiments, the cells contain DNA or RNA encoding a human  $\alpha_1$  subunit, preferably at least an  $\alpha_{1D}$  or  $\alpha_{1B}$  subunit. In more preferred embodiments, the cells contain DNA or RNA encoding additional heterologous 20 subunits, including at least one  $\beta$ ,  $\alpha_2$  or  $\gamma$  subunits are included. In such embodiments, eukaryotic cells stably or transiently transfected with any combination of one, two, three or four of the subunit-encoding cDNA clones, such as  $\alpha_1$ ,  $\alpha_1 + \beta$ ,  $\alpha_1 + \beta + \alpha_2$ , are provided. In more preferred 25 embodiments, the subunits encoded by the heterologous DNA are human subunits.

In preferred embodiments, the cells express such heterologous calcium channel subunits and include one or more of the subunits in membrane spanning heterologous 30 calcium channels. In more preferred embodiments, the eukaryotic cells express functional, heterologous calcium channels that are capable of gating the passage of calcium channel selective ions and/or binding compounds that, at physiological concentrations, modulate the activity of the 35 heterologous calcium channel. In certain embodiments, the

-9-

heterologous calcium channels include at least one heterologous calcium channel subunit. In most preferred embodiments, the calcium channels that are expressed on the surface of the eukaryotic cells are composed substantially 5 or entirely of subunits encoded by the heterologous DNA or RNA. In preferred embodiments, the heterologous calcium channels of such cells are distinguishable from any endogenous calcium channels of the host cell.

In certain embodiments the recombinant eukaryotic 10 cells that contain the heterologous DNA encoding the calcium channel subunits are produced by transfection with DNA encoding one or more of the subunits or are injected with RNA transcripts of cDNA encoding one or more of the calcium channel subunits. The DNA may be introduced as a 15 linear DNA fragment or may be included in an expression vector for stable or transient expression of the subunit-encoding DNA. Vectors containing DNA encoding human calcium channel subunits are also provided.

The eukaryotic cells that express heterologous calcium 20 channels may be used in assays for calcium channel function or, in the case of cells transformed with fewer subunit-encoding nucleic acids than necessary to constitute a functional recombinant human calcium channel, such cells may be used to assess the effects of additional subunits on 25 calcium channel activity. The additional subunits can be provided by subsequently transfecting such a cell with one or more DNA clones or RNA transcripts encoding human calcium channel subunits.

The recombinant eukaryotic cells that express membrane 30 spanning heterologous calcium channels may be used in methods for identifying compounds that modulate calcium channel activity. In particular, the cells are used in assays that identify agonists and antagonists of calcium channel activity in humans and/or assessing the

-10-

contribution of the various calcium channel subunits to the transport and regulation of transport of calcium ions.

Assays using the eukaryotic cells for identifying compounds that modulate calcium channel activity are  
5 provided.

Isolated and purified DNA fragments that encode human calcium channel subunits are provided. DNA encoding  $\alpha_1$  subunits of a human calcium channel, and RNA, encoding such subunits, made upon transcription of such DNA are provided.

10 In particular, DNA fragments encoding  $\alpha_1$  subunits of voltage-dependent human calcium channels (VDCCs) type A, type B (also referred to as VDCC IV), type C (also referred to as VDCC II) and type D (also referred to as VDCC III) are provided.

15 In particular, DNA encoding an  $\alpha_{1D}$  subunit that includes the amino acids substantially as set forth as residues 10-2161 of sequence ID No. 1 is provided.

DNA encoding an  $\alpha_{1D}$  subunit includes substantially the amino acids set forth as amino acids 1-34 in sequence ID No. 2 in  
20 place of amino acids 373-406 of SEQ ID No. 1 is also provided. DNA encoding an  $\alpha_{1C}$  subunit that includes the amino acids substantially as set forth in sequence ID No. 3 or sequence ID No. 6 and DNA encoding an  $\alpha_{1B}$  subunit that includes an amino acid sequence substantially as set forth  
25 in sequence ID No. 7 or in sequence ID No. 8 is also provided. A phage lysate of an *E. coli* host containing DNA encoding  $\alpha_{1A}$  have been deposited in the American Type Culture Collection, 12301 Parklawn Drive, Rockville, Maryland 20852 U.S.A. under Accession No. in accord  
30 with the Budapest Treaty. The DNA in such phage includes a DNA fragment having the sequence set forth in SEQ ID No.  
21. This fragment hybridizes to DNA encoding  $\alpha_{1A}$  but not to DNA encoding  $\alpha_{1B}$ .

DNA encoding  $\alpha_2$  subunits of a human calcium channel,  
35 and RNA encoding such subunits, made upon transcription of

-11-

such a DNA are provided. DNA encoding splice variants of the  $\alpha_2$  subunit, including tissue specific splice variants, are also provided. In particular, DNA encoding the  $\alpha_2$ - $\alpha_2$  subunit subtypes is provided. In particularly preferred 5 embodiments, the DNA encoding the  $\alpha_2$  subunit is produced by alternative processing of a primary transcript that includes DNA encoding the amino acids set forth in SEQ ID 11 and the DNA of SEQ ID No. 13 inserted between nucleotides 1624 and 1625 of SEQ ID No. 11.

10 Isolated and purified DNA fragments encoding human calcium channel  $\beta$  subunits, including DNA encoding  $\beta_1$  subunit splice variants and the  $\beta_3$  subunit and is provided. In particular, DNA encoding the  $\beta_1$  and  $\beta_3$  subunits, including the  $\beta_1$  subunit splice variants  $\beta_{1.1}$ - $\beta_{1.5}$ , is 15 provided. RNA, encoding  $\beta$  subunits, made upon transcription of the DNA is also provided. *Escherichia coli* (*E. coli*) containing plasmids containing DNA encoding  $\beta_3$  have been deposited in accord with the Budapest Treaty under Accession No. 69048 at the American Type Culture 20 Collection. A partial sequence of the deposited clone is set forth in SEQ ID No. 19 (sequence from the 5' end) and SEQ ID No. 20 (sequence from the 3' end).

DNA encoding  $\beta$  subunits that are produced by alternative processing of a primary transcript encoding a 25  $\beta$  subunit, including a transcript that includes DNA encoding the amino acids set forth in SEQ ID No. 9 or including a primary transcript that encodes  $\beta_3$  as deposited under ATCC Accession No. 69048, but lacking and including alternative exons are provided or may be constructed from 30 the DNA provided herein. For example, DNA encoding a  $\beta$  subunit that is produced by alternative processing of a primary transcript that includes DNA encoding the amino acids set forth in SEQ ID No. 9, but including the DNA set forth in SEQ ID No. 12 inserted in place of nucleotides 35 615-781 of SEQ ID No. 9 is also provided. DNA encoding  $\beta$

-12-

subunits that are encoded by transcripts that have the sequence set forth in SEQ ID No. 9 including the DNA set forth in SEQ ID No. 12 inserted in place of nucleotides 615-781 of SEQ ID No. 9, but that lack one or more of the 5 following sequences of nucleotides: nucleotides 14-34 of SEQ ID No. 12, nucleotides 13-34 of SEQ ID No. 12, nucleotides 35-55 of SEQ ID No 12, nucleotides 56-190 of SEQ ID No. 12 and nucleotides 191-271 of SEQ ID No. 12 are also provided.

10 DNA encoding  $\gamma$  subunits of human calcium channels is also provided. RNA, encoding  $\gamma$  subunits, made upon transcription of the DNA are also provided. In particular, DNA containing the sequence of nucleotides set forth in SEQ ID No. 14 is provided.

15 Full-length DNA clones and corresponding RNA transcripts, encoding the  $\alpha_1$ , including  $\alpha_{1D}$ ,  $\alpha_{1B}$ ,  $\alpha_2$  and  $\beta$  subunits, including  $\beta_{1-1}-\beta_{1-5}$ , of human calcium channels are provided. Also provided are DNA clones encoding substantial portions of the  $\alpha_{1A}$ ,  $\alpha_{1C}$ ,  $\beta_3$  and  $\gamma$  subunits f 20 voltage-dependent human calcium channels for the preparation of full-length DNA clones encoding the full-length  $\alpha_{1A}$ ,  $\alpha_{1C}$ ,  $\beta_3$  and  $\gamma$  subunits.

Nucleic acid probes containing at least about 14 contiguous nucleotides of  $\alpha_{1D}$ ,  $\alpha_{1C}$ ,  $\alpha_{1B}$ ,  $\alpha_{1A}$ ,  $\alpha_2$ ,  $\beta$ , including 25  $\beta_1$  splice variants and  $\beta_3$ , and  $\gamma$  subunit-encoding DNA are provided. Methods using the probes for the isolation and cloning of calcium channel subunit-encoding cDNA, including splice variants within tissues and inter-tissue variants are also provided.

30 Purified human calcium channel subunits and purified human calcium channels are provided. The subunits and channels can be isolated from a eukaryotic cell transfected with DNA that encodes the subunit.

In another embodiment, immunoglobulins or antibodies 35 obtained from the serum of an animal immunized with a

-13-

substantially pure preparation of a human calcium channel, human calcium channel subunit or epitope-containing fragment of a human calcium subunit are provided. Monoclonal antibodies produced using a human calcium 5 channel, human calcium channel subunit or epitope-containing fragment thereof as an immunogen are also provided. *E. coli* fusion proteins including a fragment of a human calcium channel subunit may also be used as immunogen. Such fusion proteins may contain a bacterial 10 protein or portion thereof, such as the *E. coli* TrpE protein, fused to a calcium channel subunit peptide. The immunoglobulins that are produced using the calcium channel subunits or purified calcium channels as immunogens have, among other properties, the ability to specifically and 15 preferentially bind to and/or cause the immunoprecipitation of a human calcium channel or a subunit thereof which may be present in a biological sample or a solution derived from such a biological sample.

A diagnostic method for determining the presence of 20 Lambert Eaton Syndrome (LES) in a human based on immunological reactivity of LES immunoglobulin G (IgG) with a human calcium channel subunit or a eukaryotic cell which expresses a recombinant human calcium channel or a subunit thereof is also provided. In particular, an immunoassay 25 method for diagnosing Lambert-Eaton Syndrome in a person by combining serum or an IgG fraction from the person (test serum) with calcium channel proteins, including the  $\alpha$  and  $\beta$  subunits, and ascertaining whether antibodies in the test serum react with one or more of the subunits, or a 30 recombinant cell which expresses one or more of the subunits to a greater extent than antibodies in control serum, obtained from a person or group of persons known to be free of the Syndrome, is provided. Any immunoassay procedure known in the art for detecting antibodies against 35 a given antigen in serum can be employed in the method.

-14-

**DETAILED DESCRIPTION OF THE INVENTION**

**Definitions:**

Unless defined otherwise, all technical and scientific terms used herein have the same meaning as is commonly understood by one of skill in the art to which this invention belongs. All patents and publications referred to herein are incorporated by reference herein.

Reference to each of the calcium channel subunits includes the subunits that are specifically disclosed herein and human calcium channel subunits encoded by DNA that can be isolated by using the DNA disclosed as probes and screening an appropriate human cDNA or genomic library under at least low stringency. Such DNA also includes DNA that encodes proteins that have about 40% homology to any of the subunits proteins described herein or DNA that hybridizes under conditions of at least low stringency to the DNA provided herein and the protein encoded by such DNA exhibits additional identifying characteristics, such as function or molecular weight.

It is understood that subunits that are encoded by transcripts that represent splice variants of the disclosed subunits or other such subunits may exhibit less than 40% overall homology to any single subunit, but will include regions of such homology to one or more such subunits. It is also understood that 40% homology refers to proteins that share approximately 40% of their amino acids in common or that share somewhat less, but include conservative amino acid substitutions, whereby the activity of the protein is not substantially altered.

As used herein, the  $\alpha_1$  subunits types, encoded by different genes, are designated as type  $\alpha_{1A}$ ,  $\alpha_{1B}$ ,  $\alpha_{1C}$ , and  $\alpha_{1D}$ . These types may also be also referred to as VDCC IV for  $\alpha_{1B}$ , VDCC II for  $\alpha_{1C}$  and VDCC III for  $\alpha_{1D}$ . Subunit subtypes, which are splice variants, are referred to, for example as  $\alpha_{1B-1}$ ,  $\alpha_{1B-2}$ ,  $\alpha_{1C-1}$  etc.

-15-

Thus, as used herein, DNA encoding the  $\alpha_1$  subunit refers to DNA that hybridizes to the DNA provided herein under conditions of at least low stringency or encodes a subunit that has roughly about 40% homology to protein 5 encoded by DNA disclosed herein that encodes an  $\alpha_1$  subunit of a human calcium. An  $\alpha_1$  subunit may be identified by its ability to form a calcium channel. Typically,  $\alpha_1$  subunits have molecular weights greater than at least about 120 kD. The activity of a calcium channel may be assessed *in vitro* 10 by methods known to those of skill in the art, including the electrophysiological and other methods described herein. Typically,  $\alpha_1$  subunits include regions to which one or more modulators of calcium channel activity, such as a 1,4 DHP or  $\omega$ -CgTx, interact directly or indirectly. Types 15 of  $\alpha_1$  subunits may be distinguished by any method known to those of skill in the art, including on the basis of binding specificity. For example, it has been found herein that  $\alpha_{1B}$  subunits participate in the formation N-type channels,  $\alpha_{1D}$  subunits participate in the formation of L- 20 type channels, and  $\alpha_{1A}$  subunits appear to participate in the formation of channels that exhibit characteristics typical of P-type channels. Thus, for example, the activity of channels that contain the  $\alpha_{1B}$  subunit are insensitive to 1,4 DHPs; whereas the activity of channels that contain the  $\alpha_{1D}$  25 subunit are modulated or altered by a 1,4 DHP. Types and subtypes of  $\alpha_1$  subunits may be characterized on the basis of the effects of such modulators on the subunit or a channel containing the subunit as well as differences in currents and current kinetics produced by calcium channels 30 containing the subunit.

As used herein, an  $\alpha_2$  subunit is encoded by DNA that hybridizes to the DNA provided herein under conditions of low stringency or encodes a protein that has about 40% homology with that disclosed herein. Such DNA encodes a 35 protein that typically has a molecular weight greater than

-16-

about 120 kD, but does not form a calcium channel in the absence of an  $\alpha_1$  subunit, and may alter the activity of a calcium channel that contains an  $\alpha_1$  subunit. Subtypes of the  $\alpha_2$  subunit that arise as splice variants are designated by lower case letter, such as  $\alpha_{2a}$ , . . .  $\alpha_{2e}$ . In addition, the  $\alpha_2$  subunit and the large fragment produced under reducing conditions appear to be glycosylated with at least N-linked sugars and do not specifically bind to th 1,4-DHPs and phenylalkylamines that specifically bind to the  $\alpha_1$  subunit. The smaller fragment, the C-terminal fragment, is referred to as the  $\delta$  subunit and includes amino acids from about 946 (SEQ ID No. 11) through about the C-terminus. This fragment may dissociate from the remaining portion of  $\alpha_2$  when the  $\alpha_2$  subunit is exposed to reducing conditions.

As used herein, a  $\beta$  subunit is encoded by DNA that hybridizes to the DNA provided herein under conditions of low stringency or encodes a protein that has about 40% homology with that disclosed herein and is a protein that typically has a molecular weight lower than the  $\alpha$  subunits and on the order of about 50-80 kD, does not form a detectable calcium channel in the absence of an  $\alpha_1$  subunit, but may alter the activity of a calcium channel that contains an  $\alpha_1$  subunit or that contains an  $\alpha_1$  and  $\alpha_2$  subunit.

Types of the  $\beta$  subunit that are encoded by different genes are designated with subscripts, such as  $\beta_1$  and  $\beta_3$ . Subtypes of  $\beta$  subunits that arise as splice variants of a particular type are designated with a numerical subscript referring to the subtype and to the variant. Such subtypes include, but are not limited to the  $\beta_1$  splice variants, including  $\beta_{1.1}-\beta_{1.5}$ .

As used herein, a  $\gamma$  subunit is a subunit encoded by DNA disclosed herein as encoding the  $\gamma$  subunit and may be isolated and identified using the DNA disclosed herein as a probe by hybridization or other such method known to

-17-

those of skill in the art, whereby full-length clones encoding a  $\gamma$  subunit may be isolated or constructed. A  $\gamma$  subunit will be encoded by DNA that hybridizes to the DNA provided herein under conditions of low stringency or  
5 exhibits sufficient sequence homology to encode a protein that has about 40% homology with the  $\gamma$  subunit described herein.

Thus, one of skill in the art, in light of the disclosure herein, can identify DNA encoding  $\alpha_1$ ,  $\alpha_2$ ,  $\beta$ ,  $\delta$  and calcium channel subunits, including types encoded by different genes and subtypes that represent splice variants. For example, DNA probes based on the DNA disclosed herein may be used to screen an appropriate library, including a genomic or cDNA library, and obtain  
10 DNA in one or more clones that includes an open reading fragment that encodes an entire protein. Subsequent to screening an appropriate library with the DNA disclosed herein, the isolated DNA can be examined for the presence of an open reading frame from which the sequence of the  
15 encoded protein may be deduced. Determination of the molecular weight and comparison with the sequences herein should reveal the identity of the subunit as an  $\alpha_1$ ,  $\alpha_2$  etc. subunit. Functional assays may, if necessary, be used to determine whether the subunit is an  $\alpha_1$ ,  $\alpha_2$  subunit or  $\beta$   
20 subunit.  
25

For example, DNA encoding  $\alpha_{1A}$  may be isolated by screening an appropriate library with DNA, encoding all or a portion of the human  $\alpha_{1A}$  subunit, isolated from the phage deposited under ATCC Accession No. , including  
30 screening with an oligonucleotide having the sequence set forth in SEQ ID No. 21. Similarly, DNA encoding  $\beta$ , may be isolated by screening a human cDNA library with DNA probes prepared from the plasmid  $\beta 1.42$  deposited under ATCC Accession No. 69048 or probes having sequences prepared  
35 according to the sequences set forth in SEQ ID Nos. 19 and

-18-

20. Any method known to those of skill in the art for isolation and identification of DNA and preparation of full-length genomic or cDNA clones, including methods exemplified herein, may be used.

5       The subunit encoded by isolated DNA may be identified by comparison with the DNA and amino acid sequences of the subunits provided herein. Splice variants share extensive regions of homology, but include non-homologous regions, subunits encoded by different genes share a uniform  
10 distribution of non-homologous sequences.

As used herein, a splice variant refers to a variant produced by differential processing of a primary transcript of genomic DNA that results in more than one type of mRNA. Splice variants may occur within a single tissue type or  
15 among tissues (tissue-specific variants). Thus, cDNA clones that encode calcium channel subunit subtypes that have regions of identical amino acids and regions of different amino acid sequences are referred to herein as "splice variants".

20      As used herein, a "calcium channel selective ion" is an ion that is capable of flowing through, or being blocked from flowing through, a calcium channel which spans a cellular membrane under conditions which would substantially similarly permit or block the flow of  $\text{Ca}^{2+}$ .  
25       $\text{Ba}^{2+}$  is an example of an ion which is a calcium channel selective ion.

As used herein, a compound that modulates calcium channel activity is one that affects the ability of the calcium channel to pass calcium channel selective ions or  
30 affects other detectable calcium channel features, such as current kinetics. Such compounds include calcium channel antagonists and agonists and compounds that exert their effect on the activity of the calcium channel directly or indirectly.

-19-

As used herein, a "substantially pure" subunit or protein is a subunit or protein that is sufficiently free of other polypeptide contaminants to appear homogeneous by SDS-PAGE or to be unambiguously sequenced.

- 5 As used herein, heterologous or foreign DNA and RNA are used interchangeably and refer to DNA or RNA that does not occur naturally as part of the genome in which it is present or which is found in a location or locations in the genome that differ from that in which it occurs in nature.
- 10 It is DNA or RNA that is not endogenous to the cell and has been artificially introduced into the cell. Examples of heterologous DNA include, but are not limited to, DNA that encodes a calcium channel subunit and DNA that encodes RNA or proteins that mediate or alter expression of endogenous
- 15 DNA by affecting transcription, translation, or other regulatable biochemical processes. The cell that expresses the heterologous DNA, such as DNA encoding the calcium channel subunit, may contain DNA encoding the same or different calcium channel subunits. The heterologous DNA
- 20 need not be expressed and may be introduced in a manner such that it is integrated into the host cell genome or is maintained episomally.

As used herein, operative linkage of heterologous DNA to regulatory and effector sequences of nucleotides, such as promoters, enhancers, transcriptional and translational stop sites, and other signal sequences, refers to the functional relationship between such DNA and such sequences of nucleotides. For example, operative linkage of heterologous DNA to a promoter refers to the physical and functional relationship between the DNA and the promoter such that the transcription of such DNA is initiated from the promoter by an RNA polymerase that specifically recognizes, binds to and transcribes the DNA in reading frame.

-20-

As used herein, isolated, substantially pure DNA refers to DNA fragments purified according to standard techniques employed by those skilled in the art [see, e.g., Maniatis et al. (1982) *Molecular Cloning: A Laboratory Manual*, Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY].

As used herein, expression refers to the process by which nucleic acid is transcribed into mRNA and translated into peptides, polypeptides, or proteins. If the nucleic acid is derived from genomic DNA, expression may, if an appropriate eukaryotic host cell or organism is selected, include splicing of the mRNA.

As used herein, vector or plasmid refers to discrete elements that are used to introduce heterologous DNA into cells for either expression of the heterologous DNA or for replication of the cloned heterologous DNA. Selection and use of such vectors and plasmids are well within the level of skill of the art.

As used herein, expression vector includes vectors capable of expressing DNA fragments that are in operative linkage with regulatory sequences, such as promoter regions, that are capable of effecting expression of such DNA fragments. Thus, an expression vector refers to a recombinant DNA or RNA construct, such as a plasmid, a phage, recombinant virus or other vector that, upon introduction into an appropriate host cell, results in expression of the cloned DNA. Appropriate expression vectors are well known to those of skill in the art and include those that are replicable in eukaryotic cells and/or prokaryotic cells and those that remain episomal or may integrate into the host cell genome.

As used herein, a promoter region refers to the portion of DNA of a gene that controls transcription of DNA to which it is operatively linked. The promoter region includes specific sequences of DNA that are sufficient for

-21-

RNA polymerase recognition, binding and transcription initiation. This portion of the promoter region is referred to as the promoter. In addition, the promoter region includes sequences that modulate this recognition, 5 binding and transcription initiation activity of the RNA polymerase. These sequences may be *cis* acting or may be responsive to *trans* acting factors. Promoters, depending upon the nature of the regulation, may be constitutive or regulated.

10 As used herein, a recombinant eukaryotic cell is a eukaryotic cell that contains heterologous DNA or RNA.

As used herein, a recombinant or heterologous calcium channel refers to a calcium channel that contains one or more subunits that are encoded by heterologous DNA that has 15 been introduced into and expressed in a eukaryotic cells that expresses the recombinant calcium channel. A recombinant calcium channel may also include subunits that are produced by DNA endogenous to the cell. In certain embodiments, the recombinant or heterologous calcium 20 channel may contain only subunits that are encoded by heterologous DNA.

As used herein, "functional" with respect to a recombinant or heterologous calcium channel means that the channel is able to provide for and regulate entry of 25 calcium channel selective ions, including, but not limited to,  $\text{Ca}^{2+}$  or  $\text{Ba}^{2+}$ , in response to a stimulus and/or bind ligands with affinity for the channel. Preferably such calcium channel activity is distinguishable, such as electrophysiological, pharmacological and other means known 30 to those of skill in the art, from any endogenous calcium channel activity that in the host cell.

As used herein, a peptide having an amino acid sequence substantially as set forth in a particular SEQ ID No. includes peptides that have the same function but may 35 include minor variations in sequence, such as conservative

-22-

amino acid changes or minor deletions or insertions that do not alter the activity of the peptide. The activity of a calcium channel receptor subunit peptide refers to its ability to form functional calcium channels with other such 5 subunits.

As used herein, a physiological concentration of a compound is that which is necessary and sufficient for a biological process to occur. For example, a physiological concentration of a calcium channel selective ion is a 10 concentration of the calcium channel selective ion necessary and sufficient to provide an inward current when the channels open.

As used herein, activity of a calcium channel refers to the movement of a calcium selective ion through a 15 calcium channel. Such activity may be measured by any method known to those of skill in the art, including, but not limited to, measurement of the amount of current which flows through the recombinant channel in response to a stimulus.

20 As used herein, a "functional assay" refers to an assay that identifies functional calcium channels. A functional assay, thus, is an assay to assess function.

As understood by those skilled in the art, assay methods for identifying compounds, such as antagonists and 25 agonists, that modulate calcium channel activity, generally requires comparison to a control. One type of a "control" cell or "control" culture is a cell or culture that is treated substantially the same as the cell or culture exposed to the test compound except that the control 30 culture is not exposed to the test compound. Another type of a "control" cell or "control" culture may be a cell or a culture of cells which are identical to the transfected cells except the cells employed for the control culture do not express functional calcium channels. In this 35 situation, the response of test cell to the test compound

-23-

compared to the response (or lack of response) of the receptor-negative cell to the test compound, when cells or cultures of each type of cell are exposed to substantially the same reaction conditions in the presence of the 5 compound being assayed. For example, in methods that use patch clamp electrophysiological procedures, the same cell can be tested in the presence and absence of the test compound, by changing the external solution bathing the cell as known in the art.

10 **Assays**

**Assays for identifying compounds that modulate calcium channel activity**

In vitro methods for identifying compounds, such as calcium channel agonist and antagonists, that modulate the 15 activity of calcium channels using eukaryotic cells that express heterologous human calcium channels are provided.

In particular, the assays use eukaryotic cells that express heterologous human calcium channel subunits encoded by heterologous DNA provided herein, for screening 20 potential calcium channel agonists and antagonists which are specific for human calcium channels and particularly for screening for compounds that are specific for particular human calcium channel subtypes. Such assays may be used in conjunction with methods of rational drug design 25 to select among agonists and antagonists, which differ slightly in structure, those particularly useful for modulating the activity of human calcium channels, and to design or select compounds that exhibit subtype- or tissue-specific calcium channel antagonist and agonist activities.

30 These assays should accurately predict the relative therapeutic efficacy of a compound for the treatment of certain disorders in humans. In addition, since subtype- and tissue-specific calcium channel subunits are provided, cells with tissue-specific or subtype-specific recombinant 35 calcium channels may be prepared and used in assays for

-24-

identification of human calcium channel tissue- or subtype-specific drugs.

The assays involve contacting the cell membrane of a recombinant eukaryotic cell which expresses at least one 5 subunit of a human calcium channel, preferably at least an  $\alpha_1$  subunit of a human calcium channel, with a test compound and measuring the ability of the test compound to specifically bind to the membrane or alter or modulate the activity of a heterologous calcium channel on the membrane.

10 In preferred embodiments, the assay uses a recombinant cell which has a calcium channel containing an  $\alpha_1$  subunit of a human calcium channel in combination with a  $\beta$ -subunit of a human calcium channel and/or an  $\alpha_2$  subunit of a human calcium channel. Recombinant cells expressing heterologous 15 calcium channels containing each of the  $\alpha_1$ ,  $\beta$  and  $\alpha_2$  human subunits, and, optionally, a  $\gamma$  subunit of a human calcium channel are especially preferred for use in such assays.

In certain embodiments, the assays for identifying compounds that modulate calcium channel activity are 20 practiced by measuring the calcium channel activity of a eukaryotic cell having a heterologous, functional calcium channel when such cell is exposed to a solution containing the test compound and a calcium channel selective ion and comparing the measured calcium channel activity to the 25 calcium channel activity of the same cell or a substantially identical control cell in a solution not containing the test compound. The cell is maintained in a solution having a concentration of calcium channel selective ions sufficient to provide an inward current when 30 the channels open. Especially preferred for use, is a recombinant cell expressing calcium channels that include each of the  $\alpha_1$ ,  $\beta$  and  $\alpha_2$  human subunits, and, optionally, a  $\gamma$  subunit of a human calcium channel. Methods for practicing such assays are known to those of skill in the 35 art. For example, for similar methods applied with *Xenopus*

-25-

*Xenopus* oocytes and acetylcholine receptors, see, Mishina et al. [(1985) *Nature* 313:364] and, with such oocytes and sodium channels [see, Noda et al. (1986) *Nature* 322:826-828]. For similar studies which have been carried 5 out with the acetylcholine receptor, see, e.g., Claudio et al. [(1987) *Science* 238:1688-1694].

The assays thus use cells, provided herein, that express heterologous functional calcium channels and measure functionally, such as electrophysiologically, the 10 ability of a test compound to potentiate, antagonize or otherwise modulate the magnitude and duration of the flow of calcium channel selective ions, such as  $\text{Ca}^{++}$  or  $\text{Ba}^{++}$ , through the heterologous functional channel. The amount of current which flows through the recombinant calcium 15 channels of a cell may be determined directly, such as electrophysiologically, or by monitoring an independent reaction which occurs intracellularly and which is directly influenced in a calcium (or other) ion dependent manner.

Any method for assessing the activity of a calcium 20 channel may be used in conjunction with the cells and assays provided herein. For example, in one embodiment of the method for testing a compound for its ability to modulate calcium channel activity, the amount of current is measured by its modulation of a reaction which is sensitiv 25 to calcium channel selective ions and uses a eukaryotic cell which expresses a heterologous calcium channel and also contains a transcriptional control element operatively linked for expression to a structural gene that encodes an indicator protein. The transcriptional control element 30 used for transcription of the indicator gene is responsive in the cell to a calcium channel selective ion, such as  $\text{Ca}^{2+}$  and  $\text{Ba}^+$ . The details of such transcriptional based assays are described in commonly owned PCT International Patent Application No. PCT/US91/5625, filed August 7, 1991, which 35 claims priority to a pending commonly owned U.S.

-26-

Application Serial No. 07/ 563,751, filed August 7, 1990, the contents of which applications are herein incorporated by reference thereto.

**Assays for diagnosis of LES**

5       LES is an autoimmune disease characterized by an insufficient release of acetylcholine from motor nerve terminals which normally are responsive to nerve impulses. Immunoglobulins (IgG) from LES patients block individual voltage-dependent calcium channels and thus inhibit calcium  
10 channel activity [Kim and Neher, *Science* 239:405-408 (1988)]. A diagnostic assay for Lambert Eaton Syndrome (LES) is provided herein. The diagnostic assay for LES relies on the immunological reactivity of LES IgG with the human calcium channels or particular subunits alone or in  
15 combination or expressed on the surface of recombinant cells. For example, such an assay may be based on immunoprecipitation of LES IgG by the human calcium channel subunits and cells that express such subunits provided herein.  
20 **Identification and isolation of DNA encoding human calcium channel subunits**

Methods for identifying and isolating DNA encoding  $\alpha_1$ ,  $\alpha_2$ ,  $\beta$  and  $\gamma$  subunits of human calcium channels are provided.

Identification and isolation of such DNA may be  
25 accomplished by hybridizing, under appropriate conditions, at least low stringency whereby DNA that encodes the desired subunit is isolated, restriction enzyme-digested human DNA with a labeled probe having at least 14 nucleotides and derived from any contiguous portion of DNA  
30 having a sequence of nucleotides set forth herein by sequence identification number. Once a hybridizing fragment is identified in the hybridization reaction, it can be cloned employing standard cloning techniques known to those of skill in the art. Full-length clones may be  
35 identified by the presence of a complete open reading frame and the identity of the encoded protein verified by

-27-

sequence comparison with the subunits provided herein and by functional assays to assess calcium channel forming ability or other function. This method can be used to identify genomic DNA encoding the subunit or cDNA encoding 5 splice variants of human calcium channel subunits generated by alternative splicing of the primary transcript of genomic subunit DNA. For instance, DNA, cDNA or genomic DNA, encoding a calcium channel subunit may be identified by hybridization to a DNA probe and characterized by 10 methods known to those of skill in the art, such as restriction mapping and DNA sequencing, and compared to the DNA provided herein in order to identify heterogeneity or divergence in the sequences the DNA. Such sequence differences may indicate that the transcripts from which 15 the cDNA was produced result from alternative splicing of a primary transcript, if the non-homologous and homologous regions are clustered, or from a different gene if the non-homologous regions are distributed throughout the cloned DNA.

20 Any suitable method for isolating genes using the DNA provided herein may be used. For example, oligonucleotides corresponding to regions of sequence differences have been used to isolate, by hybridization, DNA encoding the full-length splice variant and can be used to isolate genomic 25 clones. A probe, based on a nucleotide sequence disclosed herein, which encodes at least a portion of a subunit of a human calcium channel, such as a tissue-specific exon, may be used as a probe to clone related DNA, to clone a full-length cDNA clone or genomic clone encoding the human 30 calcium channel subunit.

Labeled, including, but not limited to, radioactively or enzymatically labeled, RNA or single-stranded DNA of at least 14 substantially contiguous bases, preferably at least 30 contiguous bases of a nucleic acid which encodes 35 at least a portion of a human calcium channel subunit, the

-28-

sequence of which nucleic acid corresponds to a segment of a nucleic acid sequence disclosed herein by reference to a SEQ ID No. are provided. Such nucleic acid segments may be used as probes in the methods provided herein for 5 cloning DNA encoding calcium channel subunits. See, generally, Sambrook et al. (1989) *Molecular Cloning: A Laboratory Manual*, 2nd Edition, Cold Spring Harbor Laboratory Press.

In addition, nucleic acid amplification techniques, 10 which are well known in the art, can be used to locate splice variants of calcium channel subunits by employing oligonucleotides based on DNA sequences surrounding the divergent sequence primers for amplifying human RNA or genomic DNA. Size and sequence determinations of the 15 amplification products can reveal splice variants. Furthermore, isolation of human genomic DNA sequences by hybridization can yield DNA containing multiple exons, separated by introns, that correspond to different splice variants of transcripts encoding human calcium channel 20 subunits.

DNA encoding types and subtypes of each of the  $\alpha_1$ ,  $\alpha_2$ ,  $\beta$  and  $\gamma$  subunit of voltage-dependent human calcium channels has been cloned herein by screening human cDNA libraries prepared from isolated poly A+ mRNA from cell lines 25 or tissue of human origin having such calcium channels. Among the sources of such cells or tissue for obtaining mRNA are human brain tissue or a human cell line of neural origin, such as a neuroblastoma cell line, human skeletal muscle or smooth muscle cells, and the like. Methods of preparing 30 cDNA libraries are well known in the art [see generally Ausubel et al. (1987) *Current Protocols in Molecular Biology*, Wiley-Interscience, New York; and Davis et al. (1986) *Basic Methods in Molecular Biology*, Elsevier Science Publishing Co., New York].

-29-

With respect to each of the respective subunits of a human calcium channel ( $\alpha_1$ ,  $\alpha_2$ ,  $\beta$  or  $\gamma$ ), once the DNA encoding the channel subunit was identified by a nucleic acid screening method, the isolated clone was used for further screening to identify overlapping clones. Some of the cloned DNA fragments can and have been subcloned into an appropriate vector such as pIBI24/25 (IBI, New Haven, CT), M13mp18/19, pGEM4, pGEM3, pGEM7Z, pSP72 and other such vectors known to those of skill in this art, and characterized by DNA sequencing and restriction enzyme mapping. A sequential series of overlapping clones may thus be generated for each of the subunits until a full-length clone can be prepared by methods, known to those of skill in the art, that include identification of translation initiation (start) and translation termination (stop) codons. For expression of the cloned DNA, the 5' noncoding region and other transcriptional and translational control regions of such a clone may be replaced with an efficient ribosome binding site and other regulatory regions as known in the art. Examples II-VI, below, describe in detail the cloning of each of the various subunits of a human calcium channel as well as subtypes and splice variants, including tissue-specific variants thereof. In the instances in which partial sequences of a subunit are disclosed, it is well within the skill of the art, in view of the teaching herein, to obtain the corresponding full-length nucleotide sequence encoding the subunit, subtype or splice variant thereof.

Identification and isolation of DNA encoding  $\alpha_1$  subunits

A number of voltage-dependent calcium channel  $\alpha_1$  subunit genes, which are expressed in the human CNS, have been identified and have been designated as  $\alpha_{1A}$ ,  $\alpha_{1B}$  (or VDCC IV),  $\alpha_{1C}$  (or VDCC II) and  $\alpha_{1D}$  (or VDCC III). DNA, isolated from a human neuronal cDNA library, that encodes each of the subunit types has been isolated. DNA encoding subtypes

-30-

of each of the types, which arise as splice variants are also provided. Subtypes are herein designated, for example, as  $\alpha_{1B-1}$ ,  $\alpha_{1B-2}$ .

The  $\alpha_1$  subunits types A, B, C, and D of voltage-dependent calcium channels, and subtypes thereof, differ with respect to sensitivity to known classes of calcium channel agonists and antagonists, such as DHPs, phenylalkylamines, omega conotoxin ( $\omega$ -CgTx) and pyrazonoylguanidines. They also appear to differ in the holding potential and ion the kinetics of currents produced upon depolarization of cell membranes containing calcium channels that include different types of  $\alpha_1$  subunits.

DNA that encodes an  $\alpha_1$ -subunit that binds to at least one compound selected from among dihydropyridines, phenylalkylamines,  $\omega$ -CgTx, components of funnel web spider toxin, and pyrazonoylguanidines is provided. For example, the  $\alpha_{1B}$  subunit provided herein appears to specifically interact with  $\omega$ -CgTx in N-type channels, and the  $\alpha_{1D}$  subunit provided herein specifically interacts with DHPs in L-type channels.

**Identification and isolation of DNA encoding the  $\alpha_{1D}$  human calcium channel subunit**

The  $\alpha_{1D}$  subunit cDNA has been isolated using fragments of the rabbit skeletal muscle calcium channel  $\alpha_1$  subunit cDNA as a probe to screen a cDNA library of a human neuroblastoma cell line, IMR32, to obtain clone  $\alpha 1.36$ . This clone was used as a probe to screen additional IMR32 cell cDNA libraries to obtain overlapping clones, which were then employed for screening until a sufficient series of clones to span the length of the nucleotide sequence encoding the human  $\alpha_{1D}$  subunit were obtained. Full-length clones encoding  $\alpha_{1D}$  were constructed by ligating portions of partial  $\alpha_{1D}$  clones as described in Example II. SEQ ID No. 1 shows the 7,635 nucleotide sequence of the cDNA encoding the  $\alpha_{1D}$  subunit. There is a 6,483 nucleotide sequence

-31-

reading frame which encodes a sequence of 2,161 amino acids (as set forth in SEQ ID No. 1).

SEQ ID No. 2 provides the sequence of an alternative exon encoding the IS6 transmembrane domain [see Tanabe, T.,  
5 et al. (1987) *Nature* 328:313-318 for a description of transmembrane domain terminology] of the  $\alpha_{1D}$  subunit.

SEQ ID No. 1 also shows the 2,161 amino acid sequence deduced from the human neuronal calcium channel  $\alpha_{1D}$  subunit DNA. Based on the amino acid sequence, the  $\alpha_{1D}$  protein has  
10 a calculated Mr of 245,163. The  $\alpha_{1D}$  subunit of the calcium channel contains four putative internal repeated sequence regions. Four internally repeated regions represent 24 putative transmembrane segments, and the amino- and carboxyl-termini extend intracellularly.

15 The  $\alpha_{1D}$  subunit has been shown to mediate DHP-sensitive, high-voltage-activated, long-lasting calcium channel activity. This calcium channel activity was detected when oocytes were co-injected with RNA transcripts encoding an  $\alpha_{1D}$  and  $\beta_1$ , or  $\alpha_{1D}$ ,  $\alpha_2$  and  $\beta_1$  subunits. This  
20 activity was distinguished from  $Ba^{2+}$  currents detected when oocytes were injected with RNA transcripts encoding the  $\beta_1$ ,  $\pm \alpha_2$  subunits. These currents pharmacologically and biophysically resembled  $Ca^{2+}$  currents reported for uninjected oocytes.

25 Identification and isolation DNA  
encoding the  $\alpha_{1A}$  human calcium channel  
subunit

Biological material containing DNA encoding the  $\alpha_{1A}$  subunit had been deposited in the American Type Culture  
30 Collection, 12301 Parklawn Drive, Rockville, Maryland 20852 U.S.A. under the terms of the Budapest Treaty on the International Recognition of Deposits of Microorganisms for Purposes of Patent Procedure and the Regulations promulgated under this Treaty. Samples of the deposited  
35 material are and will be available to industrial property offices and other persons legally entitled to receive them

-32-

under the terms of said Treaty and Regulations and otherwise in compliance with the patent laws and regulations of the United States of America and all other nations or international organizations in which this  
5 application, or an application claiming priority of this application, is filed or in which any patent granted on any such application is granted.

The  $\alpha_{1A}$  subunit is encoded by an approximately 3 kb insert in  $\lambda$ gt10 phage designated  $\alpha 1.254$  in *E. coli* host strain NM514. A phage lysate of this material has been deposited as at the American Type Culture Collection under ATCC Accession No. , as described above. DNA encoding  $\alpha_{1A}$  may also be identified by screening with a probe prepared from DNA that has SEQ ID No. 21:

15 5' CTCAGTACCATCTCTGATACCAGCCCCA 3'.

**Identification and isolation of DNA  
encoding the  $\alpha_{1B}$  human calcium channel  
subunit**

DNA encoding the  $\alpha_{1B}$  subunit was isolated by screening  
20 a human basal ganglia cDNA library with fragments of the rabbit skeletal muscle calcium channel  $\alpha_1$  subunit-encoding cDNA. A portion of one of the positive clones was used to screen an IMR32 cell cDNA library. Clones that hybridized to the basal ganglia DNA probe were used to further screen  
25 an IMR32 cell cDNA library to identify overlapping clones that in turn were used to screen a human hippocampus cDNA library. In this way, a sufficient series of clones to span nearly the entire length of the nucleotide sequence encoding the human  $\alpha_{1B}$  subunit was obtained. PCR  
30 amplification of specific regions of the IMR32 cell  $\alpha_{1B}$  mRNA yielded additional segments of the  $\alpha_{1B}$  coding sequence.

A full-length  $\alpha_{1B}$  DNA clone was constructed by ligating portions of the partial cDNA clones as described in Example II.C. SEQ ID Nos. 7 and 8 show the nucleotide sequences of  
35 DNA clones encoding the  $\alpha_{1B}$  subunit as well as the deduced amino acid sequences. The  $\alpha_{1B}$  subunit encoded by SEQ ID No.

-33-

7 is referred to as the  $\alpha_{1B-1}$  subunit to distinguish it from another  $\alpha_{1B}$  subunit,  $\alpha_{1B-2}$ , encoded by the nucleotide sequence shown as SEQ ID No. 8, which is derived from alternative splicing of the  $\alpha_{1B}$  subunit transcript.

5 PCR amplification of IMR32 cell mRNA using oligonucleotide primers designed according to nucleotide sequences within the  $\alpha_{1B-1}$ -encoding DNA has identified variants of the  $\alpha_{1B}$  transcript that appear to be splice variants because they contain divergent coding sequences.

10 Identification and isolation of DNA encoding the  $\alpha_{1C}$  human calcium channel subunit

Numerous  $\alpha_{1C}$ -specific DNA clones were isolated. Characterization of the sequence revealed the  $\alpha_{1C}$  coding sequence, the  $\alpha_{1C}$  initiation of translation sequence, and an alternatively spliced region of  $\alpha_{1C}$ . Alternatively spliced variants of the  $\alpha_{1C}$  subunit have been identified. SEQ ID No. 3 sets forth DNA encoding an  $\alpha_{1C}$  subunit. The DNA sequences set forth in SEQ ID No. 4 and No. 5 encode two possible amino terminal ends of the  $\alpha_{1C}$  protein. SEQ ID No. 6 encodes an alternative exon for the IV S3 transmembrane domain.

The isolation and identification of DNA clones encoding portions of the  $\alpha_{1C}$  subunit is described in detail 25 in Example II.

DNA encoding other  $\alpha_1$  subunits, including  $\alpha_{1A}$ , has also been isolated. Additional such subunits may also be isolated and identified using the DNA provided herein as described for the  $\alpha_{1B}$ ,  $\alpha_{1C}$  and  $\alpha_{1D}$  subunits or using other 30 methods known to those of skill in the art.

Identification and isolation DNA encoding  $\beta$  human calcium channel subunits

DNA encoding  $\beta_1$

To isolate DNA encoding the  $\beta_1$  subunit, a human 35 hippocampus cDNA library was screened by hybridization to a DNA fragment encoding a rabbit skeletal muscle calcium

-34-

channel  $\beta$  subunit. A hybridizing clone was selected and was in turn used to isolate overlapping clones until the overlapping clones encompassing DNA encoding the entire the human calcium channel  $\beta$  subunit were isolated and  
5 sequenced.

Five alternatively spliced forms of the human calcium channel  $\beta_1$  subunit have been identified and DNA encoding a number of forms have been isolated. These forms are designated  $\beta_{1.1}$ , expressed in skeletal muscle,  $\beta_{1.2}$ , expressed  
10 in the CNS,  $\beta_{1.3}$ , also expressed in the CNS,  $\beta_{1.4}$ , expressed in aorta tissue and HEK 293 cells, and  $\beta_{1.5}$ , expressed in HEK 293 cells. A full-length DNA clone encoding the  $\beta_{1.2}$  subunit has been constructed. The subunits  $\beta_{1.1}$ ,  $\beta_{1.2}$ ,  $\beta_{1.4}$  and  $\beta_{1.5}$  have been identified by PCR analysis as  
15 alternatively spliced forms of the  $\beta$  subunit.

The alternatively spliced variants were identified by comparison of amino acid sequences encoded by the human neuronal and rabbit skeletal muscle calcium channel  $\beta$  subunit-encoding DNA. This comparison revealed a 45-amino  
20 acid deletion in the human  $\beta$  subunit compared to the rabbit  $\beta$  subunit. Using DNA from the region as a probe for DNA cloning, as well as PCR analysis and DNA sequencing of this area of sequence divergence, alternatively spliced forms of the human calcium channel  $\beta$  subunit transcript were  
25 identified. For example, the sequence of DNA encoding one splice variant  $\beta_{1.2}$  is set forth in SEQ ID No. 9. SEQ ID No. 10 sets forth the sequence of the  $\beta_{1.3}$  subunit (nt 1-1851, including 3' untranslated sequence nt 1795-1851), which is another splice variant of the  $\beta$  subunit primary transcript.  
30  $\beta_{1.2}$  and  $\beta_{1.3}$  are human neuronal  $\beta$  subunits. DNA distinctive for a portion of a  $\beta$  subunit ( $\beta_{1.4}$ ) of a human aortic calcium channel and also human embryonic kidney (HEK) cells is set forth in SEQ ID No. 12 (nt 1-13 and 191-271). The sequence of DNA encoding a portion of a human calcium channel  $\beta$

-35-

subunit expressed in skeletal muscle ( $\beta_{1,1}$ ) is shown in SEQ ID No. 12 (nt 1-13 and 35-271).

DNA encoding  $\beta_3$

DNA encoding the  $\beta_3$  subunit and any splice variants

- 5 thereof may be isolated by screening a library, as described above for the  $\beta_1$  subunit, using DNA probes prepared according to SEQ ID Nos. 19 and 20 or using all or a portion of the deposited  $\beta_3$  clone plasmid  $\beta 1.42$  (ATCC Accession No. 69048).
- 10 The *E. coli* host containing plasmid  $\beta 1.42$  that includes DNA encoding the  $\beta_3$  subunit have been deposited as ATCC Accession No. 69048 in the American Type Culture Collection, 12301 Parklawn Drive, Rockville, Maryland 20852 U.S.A. under the terms of the Budapest Treaty on the
- 15 International Recognition of Deposits of Microorganisms for Purposes of Patent Procedure and the Regulations promulgated under this Treaty. Samples of the deposited material are and will be available to industrial property offices and other persons legally entitled to receive them
- 20 under the terms of said Treaty and Regulations and otherwise in compliance with the patent laws and regulations of the United States of America and all other nations or international organizations in which this application, or an application claiming priority of this
- 25 application, is filed or in which any patent granted on any such application is granted.

The  $\beta_3$  encoding plasmid is designated  $\beta 1.42$ . The plasmid contains a 2.5 kb EcoRI fragment encoding  $\beta_3$ , inserted into vector pGem<sup>®</sup>7zf(+) and has been deposited in

- 30 *E. coli* host strain DN5 $\alpha$ . A partial DNA sequence of the 5' and 3' ends of  $\beta_3$  are set forth in SEQ ID Nos. 19 and 20, respectively.

-36-

**Identification and isolation DNA encoding the  $\alpha_2$  human calcium channel subunit**

DNA encoding a human neuronal calcium channel  $\alpha_2$  subunit was isolated in a manner substantially similar to 5 that used for isolating DNA encoding an  $\alpha_1$  subunit, except that a human genomic DNA library was probed under low and high stringency conditions with a fragment of DNA encoding the rabbit skeletal muscle calcium channel  $\alpha_2$  subunit. The fragment included nucleotides having a sequence 10 corresponding to the nucleotide sequence between nucleotides 43 and 272 inclusive of rabbit back skeletal muscle calcium channel  $\alpha_2$  subunit cDNA as disclosed in PCT International Patent Application Publication No. WO 89/09834, which corresponds to U.S. Application Serial No. 15 07/620,520, which is a continuation-in-part of United States Serial No. 176,899, filed April 4, 1988, which applications have been incorporated herein by reference.

Example IV describes the isolation of DNA clones encoding  $\alpha_2$  subunits of a human calcium channel from a human 20 DNA library using genomic DNA and cDNA clones, identified by hybridization to the genomic DNA, as probes.

SEQ ID No. 11 shows the sequence of DNA encoding an  $\alpha_2$  subunit. As described in Example V, PCR analysis of RNA from human skeletal muscle, brain tissue and aorta using 25 oligonucleotide primers specific for a region of the human neuronal  $\alpha_2$  subunit cDNA that diverges from the rabbit skeletal muscle calcium channel  $\alpha_2$  subunit cDNA identified splice variants of the human calcium channel  $\alpha_2$  subunit transcript.

30 **Identification and isolation of DNA encoding  $\gamma$  human calcium channel subunits**

DNA encoding a human neuronal calcium channel  $\gamma$  subunit has been isolated as described in detail in Example VI. SEQ ID No. 14 shows the nucleotide sequence at the 3'- 35 end of this DNA which includes a reading frame encoding a sequence of 43 amino acid residues.

-37-

**Preparation of recombinant eukaryotic cells containing DNA encoding heterologous calcium channel subunits**

DNA encoding one or more of the calcium channel subunits or a portion of a calcium channel subunit may be  
5 introduced into a host cell for expression or replication of the DNA. Such DNA may be introduced using methods described in the following examples or using other procedures well known to those skilled in the art. Incorporation of cloned DNA into a suitable expression  
10 vector, transfection of eukaryotic cells with a plasmid vector or a combination of plasmid vectors, each encoding one or more distinct genes or with linear DNA, and selection of transfected cells are also well known in the art [see, e.g., Sambrook et al. (1989) *Molecular Cloning: A Laboratory Manual, Second Edition*, Cold Spring Harbor Laboratory Press].

Cloned full-length DNA encoding any of the subunits of a human calcium channel may be introduced into a plasmid vector for expression in a eukaryotic cell. Such DNA may  
20 be genomic DNA or cDNA. Host cells may be transfected with one or a combination of said plasmids, each of which encodes at least one calcium channel subunit. Alternatively, host cells may be transfected with linear DNA using methods well known to those of skill in the art.

25 While the DNA provided herein may be expressed in any eukaryotic cell, including yeast cells such as *P. pastoris* [see, e.g., Cregg et al. (1987) *Bio/Technology* 5:479], mammalian expression systems for expression of the DNA encoding the human calcium channel subunits provided herein  
30 are preferred.

The heterologous DNA may be introduced by any method known to those of skill in the art, such as transfection with a vector encoding the heterologous DNA. Particularly preferred vectors for transfection of mammalian cells are  
35 the pSV2dhfr expression vectors, which contain the SV40 early promoter, mouse dhfr gene, SV40 polyadenylation and

-38-

splice sites and sequences necessary for maintaining the vector in bacteria, cytomegalovirus (CMV) promoter-based vectors such as pCMV or pCDNA1, and MMTV promoter-based vectors. DNA encoding the human calcium channel subunits 5 has been inserted in the vector pCDNA1 at a position immediately following the CMV promoter.

Stably or transiently transfected mammalian cells may be prepared by methods known in the art by transfecting cells with an expression vector having a selectable marker 10 gene such as the gene for thymidine kinase, dihydrofolate reductase, neomycin resistance or the like, and, for transient transfection, growing the transfected cells under conditions selective for cells expressing the marker gene. Functional voltage-dependent calcium channels have been 15 produced in HEK 293 cells transfected with a derivative of the vector pCDNA1 that contains DNA encoding a human calcium channel subunit.

The heterologous DNA may be maintained in the cell as an episomal element or may be integrated into chromosomal 20 DNA of the cell. The resulting recombinant cells may then be cultured or subcultured (or passaged, in the case of mammalian cells) from such a culture or a subculture thereof. Methods for transfection, injection and culturing recombinant cells are known to the skilled artisan. 25 Eukaryotic cells in which DNA or RNA may be introduced, include any cells that are transfectable by such DNA or RNA or into which such DNA may be injected. Virtually any eukaryotic cell can serve as a vehicle for heterologous DNA. Preferred cells are those that can also express the 30 DNA and RNA and most preferred cells are those that can form recombinant or heterologous calcium channels that include one or more subunits encoded by the heterologous DNA. Such cells may be identified empirically or selected from among those known to be readily transfected or 35 injected. Preferred cells for introducing DNA include

-39-

those that can be transiently or stably transfected and include, but are not limited to cells of mammalian origin, such as COS cells, mouse L cells, CHO cells, human embryonic kidney cells, African green monkey cells and other such cells known to those of skill in the art, amphibian cells, such as *Xenopus laevis* oocytes, or those of yeast such as *Saccharomyces cerevisiae* or *Pichia pastoris*. Preferred cells for expressing injected RNA transcripts include *Xenopus laevis* oocytes. Cells that are preferred for transfection of DNA are those that can be readily and efficiently transfected. Such cells are known to those of skill in the art or may be empirically identified. Preferred cells include DG44 cells and HEK 293 cells, particularly HEK 293 cells that have been adapted for growth in suspension and that can be frozen in liquid nitrogen and then thawed and regrown. Such HEK 293 cells are described, for example in U.S. Patent No. 5,024,939 to Gorman [see, also Stillman et al. (1985) *Mol. Cell. Biol.* 5:2051-2060].

The cells may be used as vehicles for replicating heterologous DNA introduced therein or for expressing the heterologous DNA introduced therein. In certain embodiments, the cells are used as vehicles for expressing the heterologous DNA as a means to produce substantially pure human calcium channel subunits or heterologous calcium channels. Host cells containing the heterologous DNA may be cultured under conditions whereby the calcium channels are expressed. The calcium channel subunits may be purified using protein purification methods known to those of skill in the art. For example, antibodies, such as those provided herein, that specifically bind to one or more of the subunits may be used for affinity purification of the subunit or calcium channels containing the subunits. Substantially pure subunits of a human calcium channel  $\alpha_1$  subunits of a human calcium channel,  $\alpha_2$  subunits of a

-40-

human calcium channel,  $\beta$  subunits of a human calcium channel and  $\gamma$  subunits of a human calcium channel are provided. Substantially pure isolated calcium channels that contain at least one of the human calcium channel 5 subunits are also provided. Substantially pure calcium channels that contain a mixture of one or more subunits encoded by the host cell and one or more subunits encoded by heterologous DNA or RNA that has been introduced into the cell are also provided. Substantially pure subtype- r 10 tissue-type specific calcium channels are also provided.

In other embodiments, eukaryotic cells that contain heterologous DNA encoding at least one of an  $\alpha_1$  subunit of a human calcium channel, an  $\alpha_2$  subunit of a human calcium channel, a  $\beta$  subunit of a human calcium channel and a  $\gamma$  15 subunit of a human calcium channel are provided. In accordance with one preferred embodiment, the heterologous DNA is expressed in the eukaryotic cell and preferably encodes a human calcium channel  $\alpha_1$  subunit.

In particularly preferred aspects, the eukaryotic cell 20 which contains the heterologous DNA expresses it and forms a recombinant functional calcium channel activity. In more preferred aspects, the recombinant calcium channel activity is readily detectable because it is a type that is absent from the untransfected host cell or is of a magnitude not 25 exhibited in the untransfected cell.

Preferred among such cells is a recombinant eukaryotic cell with a functional heterologous calcium channel. The recombinant cell can be produced by introduction of and expression of heterologous DNA or RNA transcripts encoding 30 an  $\alpha_1$  subunit of a human calcium channel, more preferably also expressing, a heterologous DNA encoding a  $\beta$  subunit of a human calcium channel and/or heterologous DNA encoding an  $\alpha_2$  subunit of a human calcium channel. Especially preferred is the expression in such a recombinant cell of each of the 35  $\alpha_1$ ,  $\beta$  and  $\alpha_2$  subunits encoded by such heterologous DNA or

-41-

RNA transcripts, and optionally expression of heterologous DNA or an RNA transcript encoding a  $\gamma$  subunit of a human calcium channel.

In certain embodiments, the eukaryotic cell with a heterologous calcium channel is produced by introducing into the cell a first composition, which contains at least one RNA transcript that is translated in the cell into a subunit of a human calcium channel. In preferred embodiments, the subunits that are translated include an  $\alpha_1$ , 10 subunit of a human calcium channel. More preferably, the composition that is introduced contains an RNA transcript which encodes an  $\alpha_1$  subunit of a human calcium channel and also contains (1) an RNA transcript which encodes a  $\beta$  subunit of a human calcium channel and/or (2) an RNA 15 transcript which encodes an  $\alpha_2$  subunit of a human calcium channel. Especially preferred is the introduction of RNA encoding an  $\alpha_1$ , a  $\beta$  and an  $\alpha_2$  human calcium channel subunit, and, optionally, a  $\gamma$  subunit of a human calcium channel.

Methods for *in vitro* transcription of a cloned DNA and injection of the resulting RNA into eukaryotic cells are well known in the art. Transcripts of any of the full-length DNA encoding any of the subunits of a human calcium channel may be injected alone or in combination with other transcripts into eukaryotic cells for expression in the 25 cells. Amphibian oocytes are particularly preferred for expression of *in vitro* transcripts of the human calcium channel subunit cDNA clones provided herein.

The functional calcium channels may preferably include at least an  $\alpha_1$  subunit and a  $\beta$  subunit of a human calcium 30 channel. Eukaryotic cells expressing these two subunits and also cells expressing additional subunits, have been prepared by transfection of DNA and by injection of RNA transcripts. Such cells have exhibited voltage-dependent calcium channel activity attributable to calcium channels 35 that contain one or more of the heterologous human calcium

-42-

channel subunits. For example, eukaryotic cells expressing heterologous calcium channels containing an  $\alpha_2$  subunit in addition to the  $\alpha_1$  subunit and a  $\beta$  subunit have been shown to exhibit increased calcium selective ion flow across the 5 cellular membrane in response to depolarization, indicating that the  $\alpha_2$  subunit may potentiate calcium channel function.

Eukaryotic cells which express heterologous calcium channels containing at least a human  $\alpha_1$  subunit, a human  $\beta$  subunit and a human  $\alpha_2$  subunit are preferred. Eukaryotic 10 cells transformed with a composition containing cDNA or an RNA transcript that encodes an  $\alpha_1$  subunit alone or in combination with a  $\beta$  and/or an  $\alpha_2$  subunit may be used to produce cells that express functional calcium channels. Since recombinant cells expressing human calcium channels 15 containing all of the of the human subunits encoded by the heterologous cDNA or RNA are especially preferred, it is desirable to inject or transfect such host cells with a sufficient concentration of the subunit-encoding nucleic acids to form calcium channels that contain the human 20 subunits encoded by heterologous DNA or RNA. The precise amounts and ratios of DNA or RNA encoding the subunits may be empirically determined and optimized for a particular combination of subunits, cells and assay conditions.

With respect to measurement of the activity of 25 functional heterologous calcium channels, preferably, endogenous ion channel activity and, if desired, heterologous channel activity of channels that do not contain the desired subunits, of a host cell can be inhibited to a significant extent by chemical, 30 pharmacological and electrophysiological means, including the use of differential holding potential, to increase the S/N ratio of the measured heterologous calcium channel activity.

Among the uses for eukaryotic cells which 35 recombinantly express one or more subunits are assays for

-43-

determining whether a test compound has calcium channel agonist or antagonist activity. Desirably, a host cell for the expression of calcium channel subunits does not produce endogenous calcium channel subunits of the type or in an 5 amount that substantially interferes with the detection of heterologous calcium channel subunits in ligand binding assays or detection of heterologous calcium channel function, such as generation of calcium current, in functional assays.

10 With respect to ligand binding assays, the host cells preferably should not produce endogenous calcium channels which detectably interact with compounds having, at physiological concentrations (generally nanomolar or picomolar concentrations), affinity for calcium channels 15 that contain one or all of the human calcium channel subunits provided herein.

With respect to ligand binding assays for identifying a compound which has affinity for calcium channels, cells are employed which express, preferably, at least a 20 heterologous  $\alpha_1$  subunit. Transfected eukaryotic cells which express at least an  $\alpha_1$  subunit may be used to determine the ability of a test compound to specifically alter the activity of a calcium channel. Such ligand binding assays may be performed on intact transfected cells or membranes 25 prepared therefrom.

The capacity of a test compound to bind to or otherwise interact with membranes that contain heterologous calcium channels or subunits thereof may be determined by using any appropriate method, such as competitive binding 30 analysis, such as Scatchard plots, in which the binding capacity of such membranes is determined in the presence and absence of one or more concentrations of a compound having known affinity for the calcium channel. Where necessary, the results may be compared to a control 35 experiment designed in accordance with methods known to

-44-

those of skill in the art. For example, as a negative control, the results may be compared to those of assays of an identically treated membrane preparation from host cells which have not been transfected with one or more subunit-encoding nucleic acids.

5       Stably or transiently transfected cells or injected cells which express voltage-dependent human calcium channels containing one or more of the subunits of a human calcium channel desirably may be used in assays to identify  
10      agents, such as calcium channel agonists and antagonists, that modulate calcium channel activity. Functionally testing the activity of test compounds, including compounds having unknown activity, for calcium channel agonist or antagonist activity to determine if the test compound  
15      potentiates, inhibits or otherwise alters the flow of calcium through a human calcium channel can be accomplished by (a) maintaining a eukaryotic cell which is transfected or injected to express a heterologous functional calcium channel capable of regulating the flow of calcium channel  
20      selective ions into the cell in a medium containing calcium channel selective ions (i) in the presence of and (ii) in the absence of a test compound; (b) maintaining the cell under conditions such that the heterologous calcium channels are substantially closed and endogenous calcium  
25      channels of the cell are substantially inhibited (c) depolarizing the membrane of the cell maintained in step (b) to an extent and for an amount of time sufficient to cause (preferably, substantially only) the heterologous calcium channels to become permeable to the calcium channel  
30      selective ions; and (d) comparing the amount and duration of current flow into the cell in the presence of the test compound to that of the current flow into the cell, or a substantially similar cell, in the absence of the test compound.

-45-

Functional recombinant or heterologous calcium channels may be identified by any method known to those of skill in the art. For example, electrophysiological procedures for measuring the current across an ion-selective membrane of a cell, which are well known, may be used. The amount and duration of the flow of calcium selective ions through heterologous calcium channels of a recombinant cell containing DNA encoding one or more of the subunits provided herein has been measured using electrophysiological recordings using a two electrode and the whole-cell patch clamp techniques. In order to improve the sensitivity of the assays, known methods can be used to eliminate or reduce non-calcium currents and calcium currents resulting from endogenous calcium channels, when measuring calcium currents through recombinant channels. For example, the DHP Bay K 8644 specifically enhances L-type calcium channel function by increasing the duration of the open state of the channels [see, e.g., Hess, J.B., et al. (1984) *Nature* 311:538-544]. Prolonged opening of the channels results in calcium currents of increased magnitude and duration. Tail currents can be observed upon repolarization of the cell membrane after activation of ion channels by a depolarizing voltage command. The opened channels require a finite time to close or "deactivate" upon repolarization, and the current that flows through the channels during this period is referred to as a tail current. Because Bay K 8644 prolongs opening events in calcium channels, it tends to prolong these tail currents and make them more pronounced.

30

#### EXAMPLES

The following examples are included for illustrative purposes only and are not intended to limit the scope of the invention.

-46-

**EXAMPLE I: PREPARATION OF LIBRARIES USED FOR ISOLATION OF  
DNA ENCODING HUMAN NEURONAL VOLTAGE-DEPENDENT  
CALCIUM CHANNEL SUBUNITS**

**A. RNA Isolation**

5       **1. IMR32 cells**

IMR32 cells were obtained from the American Type Culture Collection (ATCC Accession No. CCL127, Rockville, MD) and grown in DMEM, 10% fetal bovine serum, 1% penicillin/streptomycin (GIBCO, Grand Island, NY) plus 1.0 10 mM dibutyryl cAMP (dbcAMP) for ten days. Total RNA was isolated from the cells according to the procedure described by H.C. Birnboim [(1988) *Nucleic Acids Research* 16:1487-1497]. Poly(A<sup>+</sup>) RNA was selected according to standard procedures [see, e.g., Sambrook et al. (1989) 15 *Molecular Cloning, A Laboratory Manual*, Cold Spring Harbor Laboratory Press; pg. 7.26-7.29].

**2. Human thalamus tissue**

Human thalamus tissue (2.34 g), obtained from the National Neurological Research Bank, Los Angeles, CA, that 20 had been stored frozen at -70°C was pulverized using a mortar and pestle in the presence of liquid nitrogen and the cells were lysed in 12 ml of lysis buffer (5 M guanidinium isothiocyanate, 50 mM TRIS, pH 7.4, 10 mM EDTA, 5% β-mercaptoethanol). Lysis buffer was added to the 25 lysate to yield a final volume of 17 ml. N-laurylsarcosine and CsCl were added to the mixture to yield final concentrations of 4% and 0.01 g/ml, respectively, in a final volume of 18 ml.

The sample was centrifuged at 9,000 rpm in a Sorvall 30 SS34 rotor for 10 min at room temperature to remove the insoluble material as a pellet. The supernatant was divided into two equal portions and each was layered onto a 2-ml cushion of a solution of 5.7 M CsCl, 0.1 M EDTA contained in separate centrifuge tubes to yield 35 approximately 9 ml per tube. The samples were centrifuged in an SW41 rotor at 37,000 rpm for 24 h at 20°C.

-47-

After centrifugation, each RNA pellet was resuspended in 3 ml ETS (10 mM TRIS, pH 7.4, 10 mM EDTA, 0.2% SDS) and combined into a single tube. The RNA was precipitated with 0.25 M NaCl and two volumes of 95% ethanol.

5       The precipitate was collected by centrifugation and resuspended in 4 ml PK buffer (0.05 M TRIS, pH 8.4, 0.14 M NaCl, 0.01 M EDTA, 1% SDS). Proteinase K was added to the sample to a final concentration of 200 µg/ml. The sample was incubated at 22°C for 1 h, followed by extraction with  
10      an equal volume of phenol:chloroform:isoamylalcohol (50:48:2) two times, followed by one extraction with an equal volume of chloroform: isoamylalcohol (24:1). The RNA was precipitated with ethanol and NaCl. The precipitate was resuspended in 400 µl of ETS buffer. The yield of  
15      total RNA was approximately 1.0 mg. Poly A<sup>+</sup> RNA (30 µg) was isolated from the total RNA according to standard methods as stated in Example I.A.1.

#### B. Library Construction

Double-stranded cDNA was synthesized according to  
20      standard methods [see, e.g., Sambrook et al. (1989) IN: *Molecular Cloning, A Laboratory Manual*, Cold Spring Harbor Laboratory Press, Chapter 8]. Each library was prepared in substantially the same manner except for differences in:  
25      1) the oligonucleotide used to prime the first strand cDNA synthesis, 2) the adapters that were attached to the double-stranded cDNA, 3) the method used to remove the free or unused adapters, and 4) the size of the fractionated cDNA ligated into the λ phage vector.

##### 1. IMR32 cDNA library #1

30      Single-stranded cDNA was synthesized using IMR32 poly(A<sup>+</sup>) RNA (Example I.A.1.) as a template and was primed using oligo (dT)<sub>12-18</sub> (Collaborative Research Inc., Bedford, MA). The single-stranded cDNA was converted to double-stranded cDNA and the yield was approximately 2µg. EcoI  
35      adapters:

-48-

5'-AATTCGGTACGTACACTCGAGC-3' = 22-mer (SEQ ID No.15)

3'- GCCATGCATGTGAGCTCG-5' = 18-mer (SEQ ID No.16)

also containing *SnaBI* and *XhoI* restriction sites were then added to the double-stranded cDNA according to the  
5 following procedure.

a. Phosphorylation of 18-mer

The 18-mer was phosphorylated using standard methods [see, e.g., Sambrook et al. (1989) IN: *Molecular Cloning, A Laboratory Manual*, Cold Spring Harbor Laboratory Press, 10 Chapter 8] by combining in a 10  $\mu$ l total volume the 18 mer (225 pmoles) with [ $^{32}$ P] $\gamma$ -ATP (7000 Ci/mmmole; 1.0  $\mu$ l) and kinase (2 U) and incubating at 37° C for 15 minutes. After incubation, 1  $\mu$ L 10 mM ATP and an additional 2 U of kinase were added and incubated at 37° C for 15 minutes.

15 Kinase was then inactivated by boiling for 10 minutes.

b. Hybridization of 22-mer

The 22-mer was hybridized to the phosphorylated 18-mer by addition of 225 pmoles of the 22-mer (plus water to bring volume to 15  $\mu$ l), and incubation at 65°C for 5 20 minutes. The reaction was then allowed to slow cool to room temperature.

The adapters were thus present at a concentration of 15 pmoles/ $\mu$ l, and were ready for cDNA-adapter ligation.

c. Ligation of adapters to cDNA

25 After the *EcoRI*, *SnaBI*, *XhoI* adapters were ligated to the double-stranded cDNA using a standard protocol [see, e.g., Sambrook et al. (1989) IN: *Molecular Cloning, A Laboratory Manual*, Cold Spring Harbor Laboratory Press, Chapter 8], the ligase was inactivated by heating the 30 mixture to 72°C for 15 minutes. The following reagents were added to the cDNA ligation reaction and heated at 37°C for 30 minutes: cDNA ligation reaction (20  $\mu$ l), water (24  $\mu$ l), 10x kinase buffer (3  $\mu$ l), 10 mM ATP (1  $\mu$ l) and kinase (2 $\mu$ l of 2 U/ $\mu$ l). The reaction was stopped by the addition

-49-

of 2  $\mu$ l 0.5M EDTA, followed by one phenol/chloroform extraction and one chloroform extraction.

d. Size Selection and Packaging of cDNA

The double-stranded cDNA with the EcoRI, SnaBI, XbaI adapters ligated was purified away from the free or unligated adapters using a 5 ml Sepharose CL-4B column (Sigma, St. Louis, MO). 100  $\mu$ l fractions were collected and those containing the cDNA, determined by monitoring the radioactivity, were pooled, ethanol precipitated, resuspended in TE buffer and loaded onto a 1% agarose gel. After the electrophoresis, the gel was stained with ethidium bromide and the 1 to 3 kb fraction was cut from the gel. The cDNA embedded in the agarose was eluted using the "Geneluter Electroporation System" (Invitrogen, San Diego, CA). The eluted cDNA was collected by ethanol precipitation and resuspended in TE buffer at 0.10 pmol/ $\mu$ l. The cDNA was ligated to 1  $\mu$ g of EcoRI digested, dephosphorylated  $\lambda$ gt11 in a 5  $\mu$ l reaction volume at a 2- to 4-fold molar excess ratio of cDNA over the  $\lambda$ gt11 vector. The ligated  $\lambda$ gt11 containing the cDNA insert was packaged into  $\lambda$  phage virions in vitro using the Gigapack (Stratagene, La Jolla, CA) kit. The packaged phage were plated on an *E. coli* Y1088 bacterial lawn in preparation for screening.

25           2. IMR32 cDNA library #2

This library was prepared as described (Example I.B.1.) with the exception that 3 to 9 kb cDNA fragments were ligated into the  $\lambda$ gt11 phage vector rather than the 1 to 3 kb fragments.

30           3. IMR32 cDNA library #3

IMR32 cell poly(A<sup>+</sup>) RNA (Example I.A.1.) was used as a template to synthesize single-stranded cDNA. The primers for the first strand cDNA synthesis were random primers (hexadeoxy-nucleotides [pd(N)<sub>6</sub>] Cat #5020-1, Clontech, Palo Alto, CA). The double-stranded cDNA was synthesized

-50-

(Example I.B.1.), EcoRI, SnaBI, XhoI adapters were added to the cDNA (Example I.B.1.), the unligated adapters were removed (Example I.B.1.), and the double-stranded cDNA with the ligated adapters was fractionated on an agarose gel 5 (Example I.B.1.). The cDNA fraction greater than 1.8 kb was eluted from the agarose (Example I.B.1.), ligated into λgt11, packaged, and plated into a bacterial lawn of Y1088 (Example I.B.1.).

#### 4. IMR32 cDNA library #4

10       IMR32 cell poly(A<sup>+</sup>) RNA (Example I.A.1.) was used as a template to synthesize single-stranded cDNA. The primers for the first strand cDNA synthesis were oligonucleotides: 89-365a specific for the  $\alpha_{1D}$  (VDCC III) type  $\alpha_1$ -subunit (see Example II.A.) coding sequence (the complementary sequence 15 of nt 2927 to 2956, SEQ ID No. 1), 89-495 specific for the  $\alpha_{1C}$  (VDCC II) type  $\alpha_1$ -subunit (see Example II.B.) coding sequence (the complementary sequence of nt 852 to 873, SEQ ID No. 3), and 90-12 specific for the  $\alpha_{1C}$ -subunit coding sequence (the complementary sequence of nt 2496 to 2520, 20 SEQ ID No. 3). The cDNA library was then constructed as described (Example I.B.3), except that the cDNA size-fraction greater than 1.5 kb was eluted from the agarose rather than the greater than 1.8 kb fraction.

#### 5. IMR32 cDNA library #5

25       The cDNA library was constructed as described (Example I.B.3.) with the exception that the size-fraction greater than 1.2 kb was eluted from the agarose rather than the greater than 1.8 kb fraction.

#### 6. Human thalamus cDNA library #6

30       Human thalamus poly (A<sup>+</sup>) RNA (Example I.A.2.) was used as a template to synthesize single-stranded cDNA. Oligo (dT) was used to prime the first strand synthesis (Example I.B.1.). The double-stranded cDNA was synthesized (Example I.B.1.) and EcoRI, KpnI, NcoI adapters of the following 35 sequence:

-51-

5' CCATGGTACCTCGTTGACG 3'= 20 mer (SEQ ID NO. 17)  
3' GGTACCATGGAAGCAACTGCTTAA 5'= 24 mer (SEQ ID NO. 18)  
were ligated to the double-stranded cDNA as described  
(Example I.B.1.) with the 20-mer replacing the 18-mer and  
5 the 24-mer replacing the 22-mer. The unligated adapters  
were removed by passing the cDNA-adapter mixture through a  
1 ml Bio Gel A-50 (Bio-Rad Laboratories, Richmond, CA.)  
column. Fractions (30  $\mu$ l) were collected and 1  $\mu$ l of each  
fraction in the first peak of radioactivity was  
10 electrophoresed on a 1% agarose gel. After  
electrophoresis, the gel was dried on a vacuum gel drier  
and exposed to x-ray film. The fractions containing cDNA  
fragments greater than 600 bp were pooled, ethanol  
precipitated, and ligated into  $\lambda$ gt11 (Example I.B.1.). The  
15 construction of the cDNA library was completed as described  
(Example I.B.1.).

**C. Hybridization and Washing Conditions**

Hybridization of radiolabelled nucleic acids to  
immobilized DNA for the purpose of screening cDNA  
20 libraries, DNA Southern transfers, or northern transfers  
was routinely performed in standard hybridization  
conditions [5 x SSPE, 5x Denhardt's, 50% deionized  
formamide, 200  $\mu$ g/ml sonicated herring sperm DNA (Cat  
#223646, Boehringer Mannheim Biochemicals, Indianapolis,  
25 IN)]. The recipes for SSPE and Denhardt's and the  
preparation of deionized formamide are described, for  
example, in Sambrook et al. (1989) *Molecular Cloning, A  
Laboratory Manual*, Cold Spring Harbor Laboratory Press,  
Chapter 8). In some hybridizations, lower stringency  
30 conditions were used in that 10% deionized formamide  
replaced 50% deionized formamide described for the standard  
hybridization conditions.

The washing conditions for removing the non-specific  
probe from the filters was either high, medium, or low  
35 stringency as described below:

-52-

- 1) high stringency: 0.1 x SSPE, 0.1% SDS, 65°C
- 2) medium stringency: 0.2 x SSPE, 0.1% SDS, 50°C
- 3) low stringency: 1.0 x SSPE, 0.1% SDS,

50°C.

5 It is understood that equivalent stringencies may be achieved using alternative buffers, salts and temperatures.

**EXAMPLE II: ISOLATION OF DNA ENCODING THE HUMAN NEURONAL CALCIUM CHANNEL  $\alpha_1$  SUBUNIT**

A. Isolation of DNA encoding the  $\alpha_{1D}$  subunit

10 1. Reference list of partial  $\alpha_{1D}$  cDNA clones

Numerous  $\alpha_{1D}$ -specific cDNA clones were isolated in order to characterize the complete  $\alpha_{1D}$  coding sequence plus portions of the 5' and 3' untranslated sequences. SEQ ID No. 1 shows the complete  $\alpha_{1D}$  DNA coding sequence, plus 510 nucleotides of  $\alpha_{1D}$  5' untranslated sequence ending in the guanidine nucleotide adjacent to the adenine nucleotide of the proposed initiation of translation as well as 642 nucleotides of 3' untranslated sequence. Also shown in SEQ ID No. 1 is the deduced amino acid sequence. A list of 15 partial cDNA clones used to characterize the  $\alpha_{1D}$  sequence and the nucleotide position of each clone relative to the full-length  $\alpha_{1D}$  cDNA sequence, which is set forth in SEQ ID No. 1, is shown below. The isolation and characterization of these clones are described below (Example II.A.2.).

25 IMR32 1.144 nt. 1 to 510 of SEQ ID No. 1  
5' untranslated sequence,

nt. 511 to 2431, SEQ ID No. 1

IMR32\* 1.136 nt. 1627 to 2988, SEQ ID No. 1  
nt. 1 to 104 of SEQ ID No. 2

30 additional exon,

IMR32@ 1.80 nt. 2083 to 6468, SEQ ID No. 1

IMR32# 1.36 nt. 2857 to 4281, SEQ ID No. 1

IMR32 1.163 nt. 5200 to 7635, SEQ ID No. 1

\* 5' of nt 1627, IMR32 1.136 encodes an intron and  
35 an additional exon described in Example II.A.2.d.

-53-

6 IMR32 1.80 contains two deletions, nt 2984 to  
3131 and nt 5303 to 5349 (SEQ ID No. 1). The  
148 nt deletion (nt. 2984 to 3131) was corrected  
by performing a polymerase chain reaction  
5 described in Example II.A.3.b.

# IMR32 1.36 contains a 132 nt deletion (nt. 3081  
to 3212).

2. Isolation and characterization of individual  
clones listed in Example II.A.1.

10 a. IMR32 1.36

Two million recombinants of the IMR32 cDNA library #1  
(Example I.B.1.) were screened in duplicate at a density of  
approximately 200,000 plaques per 150 mm plate using a  
mixture of radiolabelled fragments of the coding region of  
15 the rabbit skeletal muscle calcium channel  $\alpha_1$  cDNA [for the  
sequence of the rabbit skeletal muscle calcium channel  $\alpha_1$   
subunit cDNA, see, Tanabe et al. (1987). *Nature* 328:313-  
318]:

	Fragment	Nucleotides
	KpnI-EcoRI	-78 to 1006
20	EcoRI-XbaI	1006 to 2653
	ApaI-ApaI	3093 to 4182
	BglIII-SacI	4487 to 5310

The hybridization was performed using low stringency  
hybridization conditions (Example I.C.) and the filters  
25 were washed under low stringency (Example I.C.). Only one  
 $\alpha_{1D}$ -specific recombinant (IMR32 1.36) of the  $2 \times 10^6$  screened  
was identified. IMR32 1.36 was plaque purified by standard  
methods (J. Sambrook et al. (1989) *Molecular Cloning, A  
Laboratory Manual*, Cold Spring Harbor Laboratory Press,  
30 Chapter 8) subcloned into pGEM3 (Promega, Madison, WI) and  
characterized by DNA sequencing.

b. IMR32 1.80

Approximately  $1 \times 10^6$  recombinants of the IMR32 cDNA  
library #2 (Example I.B.2.) were screened in duplicate at  
35 a density of approximately 100,000 plaques per 150 mm plate  
using the IMR32 1.36 cDNA fragment (Example II.A.1) as a

-54-

probe. Standard hybridization conditions were used (Example I.C.), and the filters were washed under high stringency (Example I.C.). Three positive plaques were identified one of which was IMR32 1.80. IMR32 1.80 was 5 plaque purified by standard methods, restriction mapped, subcloned, and characterized by DNA sequencing.

c. IMR32 1.144

Approximately  $1 \times 10^6$  recombinants of the IMR32 cDNA library #3 (Example I.B.3) were screened with the EcoRI-PvuII fragment (nt 2083 to 2518, SEQ ID No. 1) of IMR32 10 1.80. The hybridization was performed using standard hybridization conditions (Example I.C.) and the filters were washed under high stringency (Example I.C.). Three positive plaques were identified one of which was IMR32 15 1.144. IMR32 1.144 was plaque purified, restriction mapped, and the cDNA insert was subcloned into pGEM7Z (Promega, Madison, WI) and characterized by DNA sequencing. This characterization revealed that IMR32 1.144 has a series of ATG codons encoding seven possible initiating 20 methionines (nt 511 to 531, SEQ ID No. 1). PCR analysis, and DNA sequencing of cloned PCR products encoding these seven ATG codons confirmed that this sequence is present in the  $\alpha_{1D}$  transcript expressed in dbcAMP-induced IMR32 cells.

d. IMR32 1.136

25 Approximately  $1 \times 10^6$  recombinants of the IMR32 cDNA library #4 (Example I.B.4) were screened with the EcoRI-PvuII fragment (nt 2083 to 2518, SEQ ID No. 1) of IMR32 1.80 (Example II.A.1.). The hybridization was performed using standard hybridization conditions (Example I.C.) and 30 the filters were washed under high stringency (Example I.C.). Six positive plaques were identified one of which was IMR32 1.136. IMR32 1.136 was plaque purified, restriction mapped, and the cDNA insert was subcloned into a standard plasmid vector, pSP72 (Promega, Madison, WI.), 35 and characterized by DNA sequencing. This characterization

-55-

revealed that IMR32 1.136 encodes an incompletely spliced  $\alpha_{1D}$  transcript. The clone contains nucleotides 1627 to 2988 of SEQ ID No. 1 preceded by an approximate 640 bp intron. This intron is then preceded by a 104 nt exon (SEQ ID N . 5 2) which is an alternative exon encoding the IS6 transmembrane domain [see, e.g., Tanabe et al. (1987) Nature 328:313-318 for a description of the IS1 to IVS6 transmembrane terminology] of the  $\alpha_{1D}$  subunit and can replace nt 1627 to 1730, SEQ ID No. 1, to produce a 10 completely spliced  $\alpha_{1D}$  transcript.

e. IMR32 1.163

Approximately  $1 \times 10^6$  recombinants of the IMR32 cDNA library #3 (Example I.B.3.) were screened with the NcoI-XhoI fragment of IMR32 1.80 (Example II.A.1.) containing nt 15 5811 to 6468 (SEQ ID No. 1). The hybridization was performed using standard hybridization conditions (Example I.C.) and the filters were washed under high stringency (Example I.C.). Three positive plaques were identified one of which was IMR32 1.163. IMR32 1.163 was plaque purified, 20 restriction mapped, and the cDNA insert was subcloned into a standard plasmid vector, pSP72 (Promega, Madison, WI.), and characterized by DNA sequencing. This characterization revealed that IMR32 1.163 contains the  $\alpha_{1D}$  termination codon, nt 6994 to 6996 (SEQ ID No. 1).

25 3. Construction of a full-length  $\alpha_{1D}$  cDNA [pVDCCIII(A)]

$\alpha_{1D}$  cDNA clones IMR32 1.144, IMR32 1.136, IMR32 1.80, and IMR32 1.163 (Example II.A.2.) overlap and include the entire  $\alpha_{1D}$  coding sequence, nt 511 to 6993 (SEQ ID No. 1), 30 with the exception of a 148 bp deletion, nt 2984 to 3131 (SEQ ID No. 1). Portions of these partial cDNA clones were ligated to generate a full-length  $\alpha_{1D}$  cDNA in a eukaryotic expression vector. The resulting vector was called pVDCCIII(A). The construction of pVDCCIII(A) was performed 35 in four steps described in detail below: (1) the

-56-

construction of pVDCCIII/5' using portions of IMR32 1.144, IMR32 1.136, and IMR32 1.80, (2) the construction of pVDCCIII/5'.3 that corrects the 148 nt deletion in the IMR32 1.80 portion of pVDCCIII/5', (3) the construction of 5 pVDCCIII/3'.1 using portions of IMR32 1.80 and IMR32 1.163, and (4) the ligation of a portion of the pVDCCIII/5'.3 insert, the insert of pVDCCIII/3'.1, and pcDNA1 (Invitrogen, San Diego, CA) to form pVDCCIII(A). The vector pcDNA1 is a eukaryotic expression vector containing 10 a cytomegalovirus (CMV) promoter which is a constitutive promoter recognized by mammalian host cell RNA polymerase II.

Each of the DNA fragments used in preparing the full-length construct was purified by electrophoresis through 15 an agarose gel onto DE81 filter paper (Whatman, Clifton, NJ) and elution from the filter paper using 1.0 M NaCl, 10 mM TRIS, pH 8.0, 1 mM EDTA. The ligations typically were performed in a 10  $\mu$ l reaction volume with an equal molar ratio of insert fragment and a two-fold molar excess 20 of the total insert relative to the vector. The amount of DNA used was normally about 50 ng to 100 ng.

a. pVDCCIII/5'

To construct pVDCCIII/5', IMR32 1.144 (Example II.A.2.c.) was digested with *Xho*I and *Eco*RI and the 25 fragment containing the vector (pGEM7Z),  $\alpha_{ID}$  nt 1 to 510 (SEQ ID No. 1), and  $\alpha_{ID}$  nt 511 to 1732 (SEQ ID No. 1) was isolated by gel electrophoresis. The *Eco*RI-*Apa*I fragment of IMR32 1.136 (Example II.A.2.d.) nucleotides 1732 to 2667 (SEQ ID No. 1) was isolated, and the *Apa*I-*Hind*III fragment 30 of IMR32 1.80 (Example II.A.2.b.), nucleotides 2667 to 4492 (SEQ ID No. 1) was isolated. The three DNA clones were ligated to form pVDCCIII/5' containing nt 1 to 510 (5' untranslated sequence; SEQ ID No. 1) and nt 511 to 4492 (SEQ ID No. 1).

-57-

b. pVDCCIII/5'.3

Comparison of the IMR32 1.36 and IMR32 1.80 DNA sequences revealed that these two cDNA clones differ through the  $\alpha_{1D}$  coding sequence, nucleotides 2984 to 3212.

5 PCR analysis of IMR32 1.80 and dbcAMP-induced (1.0 mM, 10 days) IMR32 cytoplasmic RNA (isolated according to Ausubel, F.M. et al. (Eds) (1988) *Current Protocols in Molecular Biology*, John Wiley and Sons, New York) revealed that IMR32 1.80 had a 148 nt deletion, nt 2984 to 3131 (SEQ ID No. 1),  
10 and that IMR32 1.36 had a 132 nt deletion, nt 3081 to 3212. To perform the PCR analysis, amplification was primed with  $\alpha_{1D}$ -specific oligonucleotides 112 (nt 2548 to 2572, SEQ ID No. 1) and 311 (the complementary sequence of nt 3928 to 3957, SEQ ID No. 1). These products were then reamplified  
15 using  $\alpha_{1D}$ -specific oligonucleotides 310 (nt 2583 to 2608 SEQ ID No. 1) and 312 (the complementary sequence of nt 3883 to 3909). This reamplified product, which contains AccI and BglII restriction sites, was digested with AccI and BglII and the AccI-BglII fragment, nt 2764 to 3890 (SEQ ID No. 1)  
20 was cloned into AccI-BglII digested pVDCCIII/5' to replace the AccI-BglII pVDCCIII/5' fragment that had the deletion. This new construct was named pVDCCIII/5'.3. DNA sequence determination of pVDCCIII/5'.3 through the amplified region confirmed the 148 nt deletion in IMR32 1.80.  
25

c. pVDCCIII/3'.1

To construct pVDCCIII/3'.1, the cDNA insert of IMR32 1.163 (Example II.A.2.e.) was subcloned into pBluescript II (Stratagene, La Jolla, CA) as an XhoI fragment. The XhoI sites on the cDNA fragment were furnished by the adapters  
30 used to construct the cDNA library (I.B.3.). The insert was oriented such that the translational orientation of the insert of IMR32 1.163 was opposite to that of the lacZ gene present in the plasmid, as confirmed by analysis of restriction enzyme digests of the resulting plasmid. This  
35 was done to preclude the possibility of expression of  $\alpha_{1D}$

-58-

sequences in DH5 $\alpha$  cells transformed with this plasmid due to fusion with the lacZ gene. This plasmid was then digested with HindIII and BglII and the HindIII - BglII fragment (the HindIII site comes from the vector and the 5 BglII site is at nt 6220, SEQ ID No. 1) was eliminated, thus deleting nt 5200 to 6220 (SEQ ID No. 1) of the IMR32 1.163 clone and removing this sequence from the remainder of the plasmid which contained the 3' BglII - XbaI fragment, nt 6220 to 7635 (SEQ ID No. 1). pVDCCIII/3'.1 10 was then made by splicing together the HindIII-PvuII fragment from IMR32 1.80 (nucleotides 4492-5294, SEQ ID No. 1), the PvuII - BglII fragment of IMR32 1.163 (nucleotides 5294 to 6220, SEQ ID No. 1) and the HindIII-BglII-digested pBluescript plasmid containing the 3' BglII/XbaI IMR32 15 1.163 fragment (nt 6220 to 7635, SEQ ID No. 1).

d. pVDCCIII(A): the full-length  $\alpha_{ID}$  construct

To construct pVDCCIII(A), the DraI-HindIII fragment (5' untranslated sequence nt 327 to 510, SEQ ID No. 1 and 20 coding sequence nt 511 to 4492, SEQ ID No. 1) of pVDCCIII/5'.3 (Example II.A.3.b.) was isolated; the HindIII-XbaI fragment of pVDCCIII/3'.1 (containing nt 4492 to 7635, SEQ ID No. 1, plus the XbaI site of the adapter) (Example II.A.3.c.) was isolated; and the plasmid vector, 25 pcDNA1, was digested with EcoRV and XbaI and isolated on an agarose gel. The three DNA fragments were ligated and MC1061-P3 (Invitrogen, San Diego, CA) was transformed. Isolated clones were analyzed by restriction mapping and DNA sequencing and pVDCCIII(A) was identified which had the 30 fragments correctly ligated together: DraI-HindIII, HindIII-XbaI, XbaI-EcoRV with the blunt-end DraI and EcoRV site ligating together to form the circular plasmid.

The amino-terminus of the  $\alpha_{ID}$  subunit is encoded by the seven consecutive 5' methionine codons (nt 511 to 531, SEQ 35 ID No. 1). This 5' portion plus nt 532 to 537,

-59-

encoding two lysine residues, were deleted from pVDCCIII(A) and replaced with an efficient ribosomal binding site (5'-ACCACC-3') to form pVDCCIII.RBS(A). Expression experiments in which transcripts of this construct were injected into 5 *Xenopus laevis* oocytes did not result in an enhancement in the recombinant voltage-dependent calcium channel expression level relative to the level of expression in oocytes injected with transcripts of pVDCCIII(A).

B. Isolation of DNA encoding the  $\alpha_{1c}$  subunit

10 1. Reference List of Partial  $\alpha_{1c}$  cDNA clones

Numerous  $\alpha_{1c}$ -specific cDNA clones were isolated in order to characterize the  $\alpha_{1c}$  coding sequence, the  $\alpha_{1c}$  initiation of translation, and an alternatively spliced region of  $\alpha_{1c}$ . SEQ ID No. 3 sets forth the characterized  $\alpha_{1c}$  15 coding sequence (nt 1 to 5904) and deduced amino acid sequence. SEQ ID No. 4 and No. 5 encode two possible amino terminal ends of the  $\alpha_{1c}$  protein. SEQ ID No. 6 encodes an alternative exon for the IV S3 transmembrane domain. Shown below is a list of clones used to characterize the  $\alpha_{1c}$  20 sequence and the nucleotide position of each clone relative to the characterized  $\alpha_{1c}$  sequence (SEQ ID No. 3). The isolation and characterization of these cDNA clones are described below (Example II.B.2).

IMR32	1.66	nt 1 to 916, SEQ ID No. 3
25		nt 1 to 132, SEQ ID No. 4
IMR32	1.157	nt 1 to 873, SEQ ID No. 3
		nt 1 to 89, SEQ ID No. 5
IMR32	1.67	nt 50 to 1717, SEQ ID No. 3
*IMR32	1.86	nt 1366 to 2583, SEQ ID No. 3
30	@1.16G	nt 758 to 867, SEQ ID No. 3
IMR32	1.37	nt 2804 to 5904, SEQ ID No. 3
CNS	1.30	nt 2199 to 3903, SEQ ID No. 3
		nt 1 to 84 of alternative exon,
		SEQ ID N . 6
35	IMR32	1.38 nt 2448 to 4702, SEQ ID No. 3

-60-

nt 1 to 84 of alternative exon,  
SEQ ID No. 6

\* IMR32 1.86 has a 73 nt deletion compared to the  
rabbit cardiac muscle calcium channel  $\alpha_1$  subunit cDNA  
5 sequence.

④ 1.16G is an  $\alpha_{1c}$  genomic clone.

2. Isolation and characterization of  
clones described in Example II.B.1.

a. CNS 1.30

10 Approximately  $1 \times 10^6$  recombinants of the human thalamus  
cDNA library No. 6 (Example I.B.6.) were screened with  
fragments of the rabbit skeletal muscle calcium channel  $\alpha_1$   
cDNA described in Example II.A.2.a. The hybridization was  
performed using standard hybridization conditions (Example  
15 I.C.) and the filters were washed under low stringency  
(Example I.C.). Six positive plaques were identified, one  
of which was CNS 1.30. CNS 1.30 was plaque purified,  
restriction mapped, subcloned, and characterized by DNA  
sequencing. CNS 1.30 encodes  $\alpha_{1c}$ -specific sequence nt 2199  
20 to 3903 (SEQ ID No. 3) followed by nt 1 to 84 of one of two  
identified alternative  $\alpha_{1c}$  exons (SEQ ID No. 6). 3' of SEQ  
ID No. 6, CNS 1.30 contains an intron and, thus, CNS 1.30  
encodes a partially spliced  $\alpha_{1c}$  transcript.

b. 1.16G

25 Approximately  $1 \times 10^6$  recombinants of a λEMBL3-based  
human genomic DNA library (Cat # HL1006d Clontech Corp.,  
Palo Alto, CA) were screened using a rabbit skeletal muscle  
cDNA fragment (nt -78 to 1006, Example II.A.2.a.). The  
hybridization was performed using standard hybridization  
30 conditions (Example I.C.) and the filters were washed under  
low stringency (Example I.C.). Fourteen positive plaques  
were identified, one of which was 1.16G. Clone 1.16G was  
plaque purified, restriction mapped, subcloned, and  
portions were characterized by DNA sequencing. DNA

-61-

sequencing revealed that 1.16G encodes  $\alpha_{1c}$ -specific sequence as described in Example II.B.1.

**c. IMR32 1.66 and IMR32 1.67**

Approximately  $1 \times 10^6$  recombinants of IMR32 cDNA library #5 (Example I.B.5.) were screened with a 151 bp *Kpn*I-*Sac*I fragment of 1.16G (Example II.B.2.b.) encoding  $\alpha_{1c}$  sequence (nt 758 to 867, SEQ ID No. 3). The hybridization was performed using standard hybridization conditions (Example I.C.). The filters were then washed in 0.5 x SSPE at 65°C. Of the positive plaques, IMR32 1.66 and IMR32 1.67 were identified. The hybridizing plaques were purified, restriction mapped, subcloned, and characterized by DNA sequencing. Two of these cDNA clones, IMR32 1.66 and 1.67, encode  $\alpha_{1c}$  subunits as described (Example II.B.1.). In addition, IMR32 1.66 encodes a partially spliced  $\alpha_{1c}$  transcript marked by a GT splice donor dinucleotide beginning at the nucleotide 3' of nt 916 (SEQ ID No. 3). The intron sequence within 1.66 is 101 nt long. IMR32 1.66 encodes the  $\alpha_{1c}$  initiation of translation, nt 1 to 3 (SEQ ID No. 3) and 132 nt of 5' untranslated sequence (SEQ ID No. 4) precede the start codon in IMR32 1.66.

**d. IMR32 1.37 and IMR32 1.38**

Approximately  $2 \times 10^6$  recombinants of IMR32 cDNA library #1 (Example I.B.1.) were screened with the CNS 1.30 CDNA fragment (Example II.B.2.a.). The hybridization was performed using low stringency hybridization conditions (Example I.C.) and the filters were washed under low stringency (Example I.C.). Four positive plaques were identified, plaque purified, restriction mapped, subcloned, and characterized by DNA sequencing. Two of the clones, IMR32 1.37 and IMR32 1.38 encode  $\alpha_{1c}$ -specific sequences as described in Example II.B.1.

DNA sequence comparison of IMR32 1.37 and IMR32 1.38 revealed that the  $\alpha_{1c}$  transcript includes two exons that encode the IVS3 transmembrane domain. IMR32 1.37 has a

-62-

single exon, nt 3904 to 3987 (SEQ ID No. 3) and IMR32 1.38 appears to be anomalously spliced to contain both exons juxtaposed, nt 3904 to 3987 (SEQ ID No. 3) followed by nt 1 to 84 (SEQ ID No. 6). The alternative splice of the  $\alpha_{1c}$  transcript to contain either of the two exons encoding the IVS3 region was confirmed by comparing the CNS 1.30 sequence to the IMR32 1.37 sequence. CNS 1.30 contains nt 1 to 84 (SEQ ID No. 6) preceded by the identical sequence contained in IMR32 1.37 for nt 2199 to 3903 (SEQ ID No. 3).

5 As described in Example II.B.2.a., an intron follows nt 1 to 84 (SEQ ID No. 6). Two alternative exons have been spliced adjacent to nt 3903 (SEQ ID No. 3) represented by CNS 1.30 and IMR32 1.37.

10

e. IMR32 1.86

15 IMR32 cDNA library #1 (Example I.B.1.) was screened in duplicate using oligonucleotide probes 90-9 (nt 1462 to 1491, SEQ ID No. 3) and 90-12 (nt 2496 to 2520, SEQ ID No. 3). These oligonucleotide probes were chosen in order to isolate a clone that encodes the  $\alpha_{1c}$  subunit between the 3' end of IMR32 1.67 (nt 1717, SEQ ID No. 3) and the 5' end of CNS 1.30 (nt 2199, SEQ ID No. 3). The hybridization conditions were standard hybridization conditions (Example I.C.) with the exception that the 50% deionized formamide was reduced to 20%. The filters were washed under low 20 stringency (Example I.C.). Three positive plaques were identified one of which was IMR32 1.86. IMR32 1.86 was plaque purified, subcloned, and characterized by restriction mapping and DNA sequencing. IMR32 1.86 encodes  $\alpha_{1c}$  sequences as described in Example II.B.1.

25 Characterization by DNA sequencing revealed that IMR32 1.86 contains a 73 nt deletion compared to the DNA encoding rabbit cardiac muscle calcium channel  $\alpha_1$  subunit [Mikami et al. (1989) *Nature* 340:230], nt 2191 to 2263. These missing nucleotides correspond to nt 2176-2248 of SEQ ID No. 3.

30 Because the 5'-end of CNS 1.30 overlaps the 3'-end of IMR32

35

-63-

1.86, some of these missing nucleotides, i.e., nt 2205-2248 of SEQ ID No. 3, are accounted for by CNS 1.30. The remaining missing nucleotides of the 73 nucleotide deletion in IMR32 1.86 (i.e., nt 2176-2204 SEQ ID No. 3) were 5 determined by PCR analysis of dbcAMP-induced IMR32 cell RNA. The 73 nt deletion is a frame-shift mutation and, thus, needs to be corrected. The exact human sequence through this region, (which has been determined by the DNA sequence of CNS 1.30 and PCR analysis of IMR32 cell RNA) 10 can be inserted into IMR32 1.86 by standard methods, e.g., replacement of a restriction fragment or site-directed mutagenesis.

f. IMR32 1.157

One million recombinants of IMR32 cDNA library #4 15 (Example I.B.4.) were screened with an *Xba*I-*Eco*RI fragment of IMR32 1.67 encoding  $\alpha_{1c}$  nt 50 to 774 (SEQ ID No. 3). The hybridization was performed using standard hybridization conditions (Example I.C.). The filters were washed under high stringency (Example I.C.). One of the positive 20 plaques identified was IMR32 1.157. This plaque was purified, the insert was restriction mapped and subcloned to a standard plasmid vector pGEM7Z (Promega, Madison, WI). The DNA was characterized by sequencing. IMR32 1.157 appears to encodes an alternative 5' portion of the  $\alpha_{1c}$  25 sequence beginning with nt 1 to 89 (SEQ ID No. 5) and followed by nt 1 to 873 (SEQ ID No. 3). Analysis of the 1.66 and 1.157 5' sequence is described below (Example II.B.3.).

30 3. Characterization of the  $\alpha_{1c}$  initiation of translation site

Portions of the sequences of IMR32 1.157 (nt 57 to 89, SEQ ID No. 5; nt 1 to 67, SEQ ID No. 3), IMR32 1.66 (nt 100 to 132, SEQ ID No. 4; nt 1 to 67, SEQ ID No. 3), were compared to the rabbit lung CaCB-receptor cDNA sequence, nt 35 -33 to 67 [Biel et al. (1990) FEBS Lett. 269:409]. The human sequences are possible alternative 5' ends of the  $\alpha_{1c}$

-64-

transcript encoding the region of initiation of translation. IMR32 1.66 closely matches the CaCB receptor cDNA sequence and diverges from the CaCB receptor cDNA sequence in the 5' direction beginning at nt 122 (SEQ ID No. 4). The start codon identified in the CaCB receptor cDNA sequence is the same start codon used to describe the  $\alpha_{1C}$  coding sequence, nt 1 to 3 (SEQ ID No. 3). The functional significance of the IMR32 1.157 sequence, nt 1 to 89 (SEQ ID No. 5), is not clear. Chimeras containing sequence between 1.157 and the  $\alpha_{1C}$  coding sequence can be constructed and functional differences can be tested.

10 C. Isolation of partial cDNA clones encoding the  $\alpha_{1B}$  subunit and construction of a full-length clone

A human basal ganglia cDNA library was screened with  
15 the rabbit skeletal muscle  $\alpha_1$  subunit cDNA fragments (see Example II.A.2.a for description of fragments) under low stringency conditions. One of the hybridizing clones was used to screen an IMR32 cell cDNA library to obtain additional partial  $\alpha_{1B}$  cDNA clones, which were in turn used  
20 to further screen an IMR32 cell cDNA library for additional partial cDNA clones. One of the partial IMR32  $\alpha_{1B}$  clones was used to screen a human hippocampus library to obtain a partial  $\alpha_{1B}$  clone encoding the 3' end of the  $\alpha_{1B}$  coding sequence. The sequence of some of the regions of the  
25 partial cDNA clones was compared to the sequence of products of PCR analysis of IMR32 cell RNA to determine the accuracy of the cDNA sequences.

PCR analysis of IMR32 cell RNA and genomic DNA using oligonucleotide primers corresponding to sequences located  
30 5' and 3' of the STOP codon of the DNA encoding the  $\alpha_{1B}$  subunit revealed an alternatively spliced  $\alpha_{1B}$ -encoding mRNA in IMR32 cells. This second mRNA product is the result of differential splicing of the  $\alpha_{1B}$  subunit transcript to include another exon that is not present in the mRNA  
35 corresponding to the other 3'  $\alpha_{1B}$  cDNA sequence that was

-65-

initially isolated. To distinguish these splice variants of the  $\alpha_{1B}$  subunit, the subunit encoded by a DNA sequence corresponding to the form containing the additional exon is referred to as  $\alpha_{1B-1}$  (SEQ ID No. 7), whereas the subunit 5 encoded by a DNA sequence corresponding to the form lacking the additional exon is referred to as  $\alpha_{1B-2}$  (SEQ ID No. 8). The sequence of  $\alpha_{1B-1}$  diverges from that of  $\alpha_{1B-2}$  beginning at nt 6633 (SEQ ID No. 7). Following the sequence of the additional exon in  $\alpha_{1B-1}$  (nt 6633-6819; SEQ ID No. 7), the  $\alpha_{1B-1}$  10 and  $\alpha_{1B-2}$  sequences are identical (i.e., nt 6820-7362 in SEQ ID No. 7 and nt 6633-7175 in SEQ ID No. 8). SEQ ID No. 7 and No. 8 set forth 143 nt of 5' untranslated sequence (nt 1-143) as well as 202 nt of 3' untranslated sequence (nt 7161-7362, SEQ ID No. 7) of the DNA encoding  $\alpha_{1B-1}$  and 321 nt 15 of 3' untranslated sequence (nt 6855-7175, SEQ ID No. 8) of the DNA encoding  $\alpha_{1B-2}$ .

PCR analysis of the IS6 region of the  $\alpha_{1B}$  transcript revealed what appear to be additional splice variants based on multiple fragment sizes seen on an ethidium bromide-stained agarose gel containing the products of the PCR 20 reaction.

A full-length  $\alpha_{1B-1}$  cDNA clone designated pcDNA- $\alpha_{1B-1}$  was prepared in an eight-step process as follows.

STEP 1: The SacI restriction site of pGEM3 (Promega, 25 Madison, WI) was destroyed by digestion at the SacI site, producing blunt ends by treatment with T4 DNA polymerase, and religation. The new vector was designated pGEMΔSac.

STEP 2: Fragment 1 (*Hind*III/*Kpn*I; nt 2337 to 4303 of SEQ 30 ID No. 7) was ligated into *Hind*III/*Kpn*I digested pGEM3ΔSac to produce p $\alpha$ 1.177HK.

STEP 3: Fragment 1 has a 2 nucleotide deletion (nt 3852 and 3853 of SEQ ID N . 7).. The deletion was repaired by inserting a PCR fragment (fragment 2) 35 of IMR32 RNA into p $\alpha$ 1.177HK. Thus, fragment 2

-66-

(*NarI/KpnI*; nt 3828 to 4303 of SEQ ID No. 7) was inserted into *NarI/KpnI* digested  $\alpha 1.177HK$  replacing the *NarI/KpnI* portion of fragment 1 and producing  $\alpha 1.177HK/PCR$ .

5 STEP 4: Fragment 3 (*KpnI/KpnI*; nt 4303 to 5663 of SEQ ID No. 7) was ligated into *KpnI* digest of  $\alpha 1.177HK/PCR$  to produce  $\alpha 1B5'K$ .

10 STEP 5: Fragment 4 (*EcoRI/HindIII*; *EcoRI* adaptor plus nt 1 to 2337 of SEQ ID No. 7) and fragment 5 (*HindIII/XhoI* fragment of  $\alpha 1B5'K$ ; nt 2337 to 5446 of SEQ ID No. 7) were ligated together into *EcoRI/XhoI* digested pcDNA1 (Invitrogen, San Diego, CA) to produce  $\alpha 1B5'$ .

15 STEP 6: Fragment 6 (*EcoRI/EcoRI*; *EcoRI* adapters on both ends plus nt 5749 to 7362 of SEQ ID No. 7) was ligated into *EcoRI* digested pBluescript II KS (Stratagene, La Jolla, CA) with the 5' end of the fragment proximal to the *KpnI* site in the polylinker to produce  $\alpha 1.230$ .

20 STEP 7: Fragment 7 (*KpnI/XhoI*; nt 4303 to 5446 of SEQ ID No. 7), and fragment 8 (*XhoI/CspI*; nt 5446 to 6259 of SEQ ID No. 7) were ligated into *KpnI/CspI* digested  $\alpha 1.230$  (removes nt 5749 to 6259 of SEQ ID No. 7 that was encoded in  $\alpha 1.230$  and maintains nt 6259 to 7362 of SEQ ID No. 7) to produce  $\alpha 1B3'$ .

25 STEP 8: Fragment 9 (*SphI/XhoI*; nt 4993 to 5446 of SEQ ID No. 7) and fragment 10 (*XhoI/XbaI* of  $\alpha 1B3'$ ; nt 5446 to 7319 of SEQ ID No. 7) were ligated into *SphI/XbaI* digested  $\alpha 1B5'$  (removes nt 4993 to 5446 of SEQ ID No. 7 that were encoded in  $\alpha 1B5'$  and maintains nt 1 to 4850 of SEQ ID No. 7) to produce pcDNA $\alpha_{1B-1}$ .

30 The resulting construct, pcDNA $\alpha_{1B-1}$ , contains, in pcDNA1, a full-length coding region encoding  $\alpha_{1B-1}$  (nt 144-

-67-

7362, SEQ ID No. 7), plus 5' untranslated sequence (nt 1-143, SEQ ID No. 7) and 3' untranslated sequence (nt 7161-7319, SEQ ID No. 7) under the transcriptional control of the CMV promoter.

5 EXAMPLE III: ISOLATION OF cDNA CLONES ENCODING THE HUMAN NEURONAL CALCIUM CHANNEL  $\beta_1$  subunit

A. Isolation of partial cDNA clones encoding the  $\beta$  subunit and construction of a full-length clone encoding the  $\beta_1$  subunit

10 A human hippocampus cDNA library was screened with the rabbit skeletal muscle calcium channel  $\beta_1$  subunit cDNA fragment (nt 441 to 1379) [for isolation and sequence of the rabbit skeletal muscle calcium channel  $\beta_1$  subunit cDNA, see U.S. Patent Application Serial NO. 482,384 or Ruth et al. (1989) *Science* 245:1115] using standard hybridization conditions (Example I.C.). A portion of one of the hybridizing clones was used to rescreen the hippocampus library to obtain additional cDNA clones. The cDNA inserts of hybridizing clones were characterized by restriction mapping and DNA sequencing and compared to the rabbit skeletal muscle calcium channel  $\beta_1$  subunit cDNA sequence.

20 Portions of the partial  $\beta_1$  subunit cDNA clones were ligated to generate a full-length clone encoding the entire  $\beta_1$  subunit. SEQ ID No. 9 shows the  $\beta_1$  subunit coding sequence (nt 1-1434) as well as a portion of the 3' untranslated sequence (nt 1435-1546). The deduced amino acid sequence is also provided in SEQ ID No. 9. In order to perform expression experiments, full-length  $\beta_1$  subunit cDNA clones were constructed as follows.

30 Step 1: DNA fragment 1 (~800 bp of 5' untranslated sequence plus nt 1-277 of SEQ ID No. 9) was ligated to DNA fragment 2 (nt 277-1546 of SEQ ID N . 9 plus 448 bp of intron sequence) and cloned into pGEM7Z. The resulting plasmid, p $\beta$ 1-1.18, contained a full-length  $\beta_1$  subunit clone 35 that included a 448-bp intron.

-68-

Step 2: To replace the 5' untranslated sequence of p $\beta$ 1-1.18 with a ribosome binding site, a double-stranded adapter was synthesized that contains an EcoRI site, sequence encoding a ribosome binding site (5'-ACCACC-3') and nt 1-25 of SEQ ID No. 9. The adapter was ligated to SmaI-digested p $\beta$ 1-1.18, and the products of the ligation reaction were digested with EcoRI.

Step 3: The EcoRI fragment from step 2 containing the EcoRI adapter, efficient ribosome binding site and nt 1-1546 of SEQ ID No. 9 plus intron sequence was cloned into a plasmid vector and designated p $\beta$ 1-1.18RBS. The EcoRI fragment of p $\beta$ 1-1.18RBS was subcloned into EcoRI-digested pcDNA1 with the initiation codon proximal to CMV promoter to form pHBCaCH $\beta$ <sub>1</sub>RBS(A).

Step 4: To generate a full-length clone encoding the  $\beta$ <sub>1</sub> subunit lacking intron sequence, DNA fragment 3 (nt 69-1146 of SEQ ID No. 9 plus 448 bp of intron sequence followed by nt 1147-1546 of SEQ ID No. 9), was subjected to site-directed mutagenesis to delete the intron sequence, thereby yielding p $\beta$ 1(-). The EcoRI-XhoI fragment of p $\beta$ 1-1.18RBS (containing of the ribosome binding site and nt 1-277 of SEQ ID No. 9) was ligated to the XhoI-EcoRI fragment of p $\beta$ 1(-) (containing of nt 277-1546 of SEQ ID No. 9) and cloned into pcDNA1 with the initiation of translation proximal to the CMV promoter. The resulting expression plasmid was designated pHBCaCH $\beta$ <sub>1b</sub>RBS(A).

#### B. Splice Variant $\beta$ <sub>13</sub>

DNA sequence analysis of the DNA clones encoding the  $\beta$ <sub>1</sub> subunit indicated that in the CNS at least two alternatively spliced forms of the same human  $\beta$ <sub>1</sub> subunit primary transcript are expressed. One form is represented by the sequence shown in SEQ ID No. 9 and is referred to as  $\beta$ <sub>1.2</sub>. The sequences of  $\beta$ <sub>1.2</sub> and the alternative form,

-69-

$\beta_{1,3}$ , diverge at nt 1334 (SEQ ID No. 9). The complete  $\beta_{1,3}$  sequence (nt 1-1851), including 3' untranslated sequence (nt 1795-1851), is set forth in SEQ ID No. 10.

5 EXAMPLE IV: ISOLATION OF cDNA CLONES ENCODING THE HUMAN NEURONAL CALCIUM CHANNEL  $\alpha_2$ -subunit

A. Isolation of cDNA clones

The complete human neuronal  $\alpha_2$  coding sequence (nt 35-3307) plus a portion of the 5' untranslated sequence (nt 1 to 34) as well as a portion of the 3' untranslated sequence 10 (nt 3308-3600) is set forth in SEQ ID No. 11.

To isolate DNA encoding the human neuronal  $\alpha_2$  subunit, human  $\alpha_2$  genomic clones first were isolated by probing human genomic Southern blots using a rabbit skeletal muscle calcium channel  $\alpha_2$  subunit cDNA fragment [nt 43 to 272, 15 Ellis et al. (1988) *Science* 240:1661]. Human genomic DNA was digested with *Eco*RI, electrophoresed, blotted, and probed with the rabbit skeletal muscle probe using standard hybridization conditions (Example I.C.) and low stringency washing conditions (Example I.C.). Two restriction 20 fragments were identified, 3.5 kb and 3.0 kb. These *Eco*RI restriction fragments were cloned by preparing a  $\lambda$ gt11 library containing human genomic *Eco*RI fragments ranging from 2.2 kb to 4.3 kb. The library was screened as described above using the rabbit  $\alpha_2$  probe, hybridizing 25 clones were isolated and characterized by DNA sequencing. HGCaCH $\alpha$ 2.20 contained the 3.5 kb fragment and HGCaCH $\alpha$ 2.9 contained the 3.0 kb fragment.

Restriction mapping and DNA sequencing revealed that HGCaCH $\alpha$ 2.20 contains an 82 bp exon (nt 130 to 211 of the 30 human  $\alpha_2$  coding sequence, SEQ ID No. 11) on a 650 bp *Pst*I-*Xba*I restriction fragment and that HGCaCH $\alpha$ 2.9 contains 105 bp of an exon (nt 212 to 316 of the coding sequence, SEQ ID No. 11) on a 750 bp *Xba*I-*Bgl*III restriction fragment. These restriction fragments were used to screen the human 35 basal ganglia cDNA library (Example II.C.2.a.). HBCaCH $\alpha$ 2.1

-70-

was isolated (nt 29 to 1163, SEQ ID No. 11) and used to screen a human brain stem cDNA library (ATCC Accession No. 37432) obtained from the American Type Culture Collection, 12301 Parklawn Drive, Rockville, MD. 20852. Two clones  
5 were isolated, HBCaCH $\alpha$ 2.5 (nt 1 to 1162, SEQ ID No. 11) and HBCaCH $\alpha$ 2.8 (nt 714 to 1562, SEQ ID No. 11, followed by 1600 nt of intervening sequence). A 2400 bp fragment of HBCaCH $\alpha$ 2.8 (beginning at nt 759 of SEQ ID No. 11 and ending at a *Sma*I site in the intron) was used to rescreen the  
10 brain stem library and to isolate HBCaCH $\alpha$ 2.11 (nt 879 to 3600, SEQ ID No. 11). Clones HBCaCH $\alpha$ 2.5 and HBCaCH $\alpha$ 2.11 overlap to encode an entire human brain  $\alpha_2$  protein.

B. Construction of pHBCaCH $\alpha$ ,A

To construct pHBCaCH $\alpha$ ,A containing DNA encoding a full-length human calcium channel  $\alpha_2$  subunit, an (*Eco*RI)-*Pvu*II fragment of HBCaCH $\alpha$ 2.5 (nt 1 to 1061, SEQ ID No. 11, *Eco*RI adapter, *Pvu*II partial digest) and a *Pvu*II-*Pst*I fragment of HBCaCH $\alpha$ 2.11 (nt 1061 to 2424 SEQ ID No. 11; *Pvu*II partial digest) were ligated into *Eco*RI-*Pst*I-digested pIBI24  
15 (Stratagene, La Jolla, CA). Subsequently, an (*Eco*RI)-*Pst*I fragment (nt 1 to 2424 SEQ ID No. 11) was isolated and ligated to a *Pst*I-(*Eco*RI) fragment (nt 2424 to 3600 SEQ ID No. 11) of HBCaCH $\alpha$ 2.11 in *Eco*RI-digested pIBI24 to produce DNA, HBCaCH $\alpha$ 2, encoding a full-length human brain  $\alpha_2$   
20 subunit. The 3600 bp *Eco*RI insert of HBCaCH $\alpha$ 2 (nt 1 to 3600, SEQ ID No. 11) was subcloned into pcDNA1 (pHBCaCH $\alpha$ 2A) with the methionine initiating codon proximal to the CMV promoter. The 3600 bp *Eco*RI insert of HBCaCH $\alpha$ 2 was also subcloned into pSV2dHFR [Subramani et al. (1981). *Mol.*  
25 *Cell. Biol.* 1:854-864] which contains the SV40 early promoter, mouse dihydrofolate reductase (dhfr) gene, SV40 polyadenylation and splice sites and sequences required for maintenance of the vector in bacteria.  
30

-71-

**EXAMPLE V. DIFFERENTIAL PROCESSING OF THE HUMAN  $\beta_1$  TRANSCRIPT AND THE HUMAN  $\alpha_2$  TRANSCRIPT**

**A. Differential processing of the  $\beta_1$  transcript**

PCR analysis of the human  $\beta_1$  transcript present in skeletal muscle, aorta, hippocampus and basal ganglia, and HEK 293 cells revealed differential processing of the region corresponding to nt 615-781 of SEQ ID No. 9 in each of the tissues. Four different sequences that result in five different processed  $\beta_1$  transcripts through this region were identified. The  $\beta_1$  transcripts from the different tissues contained different combinations of the four sequences, except for one of the  $\beta_1$  transcripts expressed in HEK 293 cells ( $\beta_{1,5}$ ) which lacked all four sequences.

None of the  $\beta_1$  transcripts contained each of the four sequences; however, for ease of reference, all four sequences are set forth end-to-end as a single long sequence in SEQ ID No. 12. The four sequences that are differentially processed are sequence 1 (nt 14-34 in SEQ ID No. 12), sequence 2 (nt 35-55 in SEQ ID No. 12), sequence 3 (nt 56-190 in SEQ ID No. 12) and sequence 4 (nt 191-271 in SEQ ID No. 12). The forms of the  $\beta_1$  transcript that have been identified include: (1) a form that lacks sequence 1 called  $\beta_{1,1}$  (expressed in skeletal muscle), (2) a form that lacks sequences 2 and 3 called  $\beta_{1,2}$  (expressed in CNS), (3) a form that lacks sequences 1, 2 and 3 called  $\beta_{1,4}$  (expressed in aorta and HEK cells) and (4) a form that lacks sequences 1-4 called  $\beta_{1,5}$  (expressed in HEK cells). Additionally, the  $\beta_{1,4}$  and  $\beta_{1,5}$  forms contain the guanine nucleotide (nt 13 in SEQ ID No. 12) which is absent in the  $\beta_{1,1}$  and  $\beta_{1,2}$  forms.

**B. Differential processing of transcripts encoding the  $\alpha_2$  subunit.**

The complete human neuronal  $\alpha_2$  coding sequence (nt 35-3307) plus a portion of the 5' untranslated sequence (nt 1 to 34) as well as a portion of the 3' untranslated sequence (nt 3308-3600) is set forth as SEQ ID No. 11.

-72-

PCR analysis of the human  $\alpha_2$  transcript present in skeletal muscle, aorta, and CNS revealed differential processing of the region corresponding to nt 1595-1942 of SEQ ID No. 11 in each of the tissues.

- 5      The analysis indicated that the primary transcript of the genomic DNA that includes the nucleotides corresponding to nt. 1595-1942 also includes an additional sequence (SEQ ID No. 13: 5' CCTATTGGTGTAGGTATAACCAACAATTAATTT AAGAAAAAGGAGACCCAATATCCAG 3') inserted between nt. 1624 and
- 10     1625 of SEQ ID No. 11. Five alternatively spliced variant transcripts that differ in the presence or absence of one to three different portions of the region of the primary transcript that includes the region of nt. 1595-1942 of SEQ ID No. 11 plus SEQ ID No. 13 inserted between nt. 1624 and
- 15     1625 have been identified. The five  $\alpha_2$ -encoding transcripts from the different tissues include different combinations of the three sequences, except for one of the  $\alpha_2$  transcripts expressed in aorta which lacks all three sequences. None of the  $\alpha_2$  transcripts contained each of the three sequences.
- 20     The sequences of the three regions that are differentially processed are sequence 1 (SEQ ID No. 13), sequence 2 ( 5' AACCCCCAAATCTCAG 3', which is nt. 1625-1639 of SEQ ID No. 11), and sequence 3 ( 5' CAAAAAAAGGGCAAAATGAAGG 3', which is nt 1908-1928 of SEQ ID No. 11). The five  $\alpha_2$  forms
- 25     identified are (1) a form that lacks sequence 3 called  $\alpha_{2a}$  (expressed in skeletal muscle), (2) a form that lacks sequence 1 called  $\alpha_{2b}$  (expressed in CNS), (3) a form that lacks sequences 1 and 2 called  $\alpha_{2c}$  expressed in aorta), (4) a form that lacks sequences 1, 2 and 3 called  $\alpha_{2d}$  (expressed in aorta) and (5) a form that lacks sequences 1 and 3 called  $\alpha_{2e}$  (expressed in aorta).

-73-

**EXAMPLE VI: ISOLATION OF DNA ENCODING A CALCIUM CHANNEL  
γ SUBUNIT FROM A HUMAN BRAIN cDNA LIBRARY**

**A. Isolation of DNA encoding the γ subunit**

Approximately  $1 \times 10^6$  recombinants from a λgt11-based 5 human hippocampus cDNA library (Clontech catalog #HL1088b, Palo Alto, CA) were screened by hybridization to a 484 bp sequence of the rabbit skeletal muscle calcium channel γ subunit cDNA (nucleotides 621-626 of the coding sequence plus 438 nucleotides of 3'-untranslated sequence) contained 10 in vector γJ10 [Jay, S. et al. (1990). *Science* 248:490-492]. Hybridization was performed using moderate stringency conditions (20% deionized formamide, 5x Denhardt's, 6 x SSPE, 0.2% SDS, 20 µg/ml herring sperm DNA, 42°C) and the filters were washed under low stringency (see 15 Example I.C.). A plaque that hybridized to this probe was purified and insert DNA was subcloned into pGEM7Z. This cDNA insert was designated γ1.4.

**B. Characterization of γ1.4**

γ1.4 was confirmed by DNA hybridization and 20 characterized by DNA sequencing. The 1500 bp *Sst*I fragment of γ1.4 hybridized to the rabbit skeletal muscle calcium channel γ subunit cDNA γJ10 on a Southern blot. SEQ analysis of this fragment revealed that it contains of approximately 500 nt of human DNA sequence and ~1000 nt of 25 λgt11 sequence (included due to apparent destruction of one of the *Eco*RI cloning sites in λgt11). The human DNA sequence contains of 129 nt of coding sequence followed immediately by a translational STOP codon and 3' untranslated sequence (SEQ ID No. 14).

To isolate the remaining 5' sequence of the human γ subunit cDNA, human CNS cDNA libraries and/or preparations of mRNA from human CNS tissues can first be assayed by PCR methods using oligonucleotide primers based on the γ cDNA-specific sequence of γ1.4. Additional human neuronal γ 35 subunit-encoding DNA can be isolated from cDNA libraries that, based on the results of the PCR assay, contain γ-specific

-74-

amplifiable cDNA. Alternatively, cDNA libraries can be constructed from mRNA preparations that, based on the results of PCR assays, contain  $\gamma$ -specific amplifiable transcripts. Such libraries are constructed by standard 5 methods using oligo dT to prime first-strand cDNA synthesis from poly A<sup>+</sup> RNA (see Example I.B.). Alternatively, first-strand cDNA can be specified by priming first-strand cDNA synthesis with a  $\gamma$  cDNA-specific oligonucleotide based on the human DNA sequence in  $\gamma$ 1.4. A cDNA library can then be 10 constructed based on this first-strand synthesis and screened with the  $\gamma$ -specific portion of  $\gamma$ 1.4.

**EXAMPLE VII: RECOMBINANT EXPRESSION OF HUMAN NEURONAL CALCIUM CHANNEL SUBUNIT-ENCODING cDNA AND RNA TRANSCRIPTS IN MAMMALIAN CELLS**

15       A. Recombinant Expression of the Human  
Neuronal Calcium Channel  $\alpha_2$  subunit cDNA  
in DG44 Cells

1. Stable transfection of DG44 cells

DG44 cells [dhfr<sup>r</sup> Chinese hamster ovary cells; see, 20 e.g., Urlaub, G. et al. (1986) *Som. Cell Molec. Genet.* 12:555-566] obtained from Lawrence Chasin at Columbia University were stably transfected by CaPO<sub>4</sub> precipitation methods [Wigler et al. (1979) *Proc. Natl. Acad. Sci. USA* 76:1373-1376] with pSV2dhfr vector containing the human 25 neuronal calcium channel  $\alpha_2$ -subunit cDNA (see Example IV) for polycistronic expression/selection in transfected cells. Transfectants were grown on 10% DMEM medium without hypoxanthine or thymidine in order to select cells that had incorporated the expression vector. Twelve transfectant 30 cell lines were established as indicated by their ability to survive on this medium.

2. Analysis of  $\alpha_2$  subunit cDNA expression in  
transfected DG44 cells

Total RNA was extracted according to the method of 35 Birnboim [(1988) *Nuc. Acids Res.* 16:1487-1497] from four of the DG44 cell lines that had been stably transfected with pSV2dhfr containing the human neuronal calcium channel  $\alpha_2$

-75-

subunit cDNA. RNA (~15 µg per lane) was separated on a 1% agarose formaldehyde gel, transferred to nitrocellulose and hybridized to the random-primed human neuronal calcium channel  $\alpha_2$  cDNA (hybridization: 50% formamide, 5 x SSPE, 5 x Denhardt's, 42° C.; wash: 0.2 x SSPE, 0.1% SDS, 65° C.). Northern blot analysis of total RNA from four of the DG44 cell lines that had been stably transfected with pSV2dhfr containing the human neuronal calcium channel  $\alpha_2$  subunit cDNA revealed that one of the four cell lines contained hybridizing mRNA the size expected for the transcript of the  $\alpha_2$  subunit cDNA (5000 nt based on the size of the cDNA) when grown in the presence of 10 mM sodium butyrate for two days. Butyrate nonspecifically induces transcription and is often used for inducing the SV40 early promoter [Gorman, C. and Howard, B. (1983) *Nucleic Acids Res.* 11:1631]. This cell line, 44 $\alpha_2$ -9, also produced mRNA species smaller (several species) and larger (6800 nt) than the size expected for the transcript of the  $\alpha_2$  cDNA (5000 nt) that hybridized to the  $\alpha_2$  cDNA-based probe. The 5000- and 6800-nt transcripts produced by this transfectant should contain the entire  $\alpha_2$  subunit coding sequence and therefore should yield a full-length  $\alpha_2$  subunit protein. A weakly hybridizing 8000-nucleotide transcript was present in untransfected and transfected DG44 cells. Apparently, DG44 cells transcribe a calcium channel  $\alpha_2$  subunit or similar gene at low levels. The level of expression of this endogenous  $\alpha_2$  subunit transcript did not appear to be affected by exposing the cells to butyrate before isolation of RNA for northern analysis.

Total protein was extracted from three of the DG44 cell lines that had been stably transfected with pSV2dhfr containing the human neuronal calcium channel  $\alpha_2$  subunit cDNA. Approximately  $10^7$  cells were sonicated in 300 µl of a solution containing 50 mM HEPES, 1 mM EDTA, 1 mM PMSF. An equal volume of 2x loading dye [Laemmli, U.K. (1970)].

-76-

Nature 227:680] was added to the samples and the protein was subjected to electrophoresis on an 8% polyacrylamide gel and then electrotransferred to nitrocellulose. The nitrocellulose was incubated with polyclonal guinea pig 5 antisera (1:200 dilution) directed against the rabbit skeletal muscle calcium channel  $\alpha_2$  subunit (obtained from K. Campbell, University of Iowa) followed by incubation with [ $^{125}$ I]-protein A. The blot was exposed to X-ray film at - 70° C. Reduced samples of protein from the transfected 10 cells as well as from untransfected DG44 cells contained immunoreactive protein of the size expected for the  $\alpha_2$  subunit of the human neuronal calcium channel (130-150 kDa). The level of this immunoreactive protein was higher 15 in 44 $\alpha_2$ -9 cells that had been grown in the presence of 10 mM sodium butyrate than in 44 $\alpha_2$ -9 cells that were grown in the absence of sodium butyrate. These data correlate well with those obtained in northern analyses of total RNA from 44 $\alpha_2$ -9 and untransfected DG44 cells. Cell line 44 $\alpha_2$ -9 also produced a 110 kD immunoreactive protein that may be either 20 a product of proteolytic degradation of the full-length  $\alpha_2$  subunit or a product of translation of one of the shorter (<5000 nt) mRNAs produced in this cell line that hybridized to the  $\alpha_2$  subunit cDNA probe.

25        B. Expression of DNA encoding human neuronal calcium channel  $\alpha_1$ ,  $\alpha_2$ , and  $\beta_1$  subunits in HEK cells

Human embryonic kidney cells (HEK 293 cells) were transiently and stably transfected with human neuronal DNA encoding calcium channel subunits. Individual 30 transfectants were analyzed electrophysiologically for the presence of voltage-activated barium currents and functional recombinant voltage-dependent calcium channels were.

1. Transfection of HEK 293 cells

35        Separate expression vectors containing DNA encoding human neuronal calcium channel  $\alpha_{1D}$ ,  $\alpha_2$  and  $\beta_1$  subunits,

-77-

plasmids pVDCCIII(A), pHBCaCH $\alpha_2$ A, and pHBCaCH $\beta_1$ RBS(A), respectively, were constructed as described in Examples II.A.3., IV.B. and III.B.3., respectively. These three vectors were used to transiently co-transfect HEK 293  
5 cells. For stable transfection of HEK 293 cells, vector pHBCaCH $\beta_1$ RBS(A) (Example III.B.3.) was used in place of pHBCaCH $\beta_1$ RBS(A) to introduce the DNA encoding the  $\beta_1$  subunit into the cells along with pVDCCIII(A) and pHBCaCH $\alpha_2$ A.

a. Transient transfection

10 Expression vectors pVDCCIII(A), pHBCaCH $\alpha_2$ A and pHBCaCH $\beta_1$ RBS(A) were used in two sets of transient transfections of HEK 293 cells (ATCC Accession No. CRL1573). In one transfection procedure, HEK 293 cells were transiently cotransfected with the  $\alpha_1$  subunit cDNA  
15 expression plasmid, the  $\alpha_1$  subunit cDNA expression plasmid, the  $\beta_1$  subunit cDNA expression plasmid and plasmid pCMV $\beta$ gal (Clontech Laboratories, Palo Alto, CA). Plasmid pCMV $\beta$ gal contains the lacZ gene (encoding *E. coli*  $\beta$ -galactosidase) fused to the cytomegalovirus (CMV) promoter and was  
20 included in this transfection as a marker gene for monitoring the efficiency of transfection. In the other transfection procedure, HEK 293 cells were transiently co-transfected with the  $\alpha_1$  subunit cDNA expression plasmid pVDCCIII(A) and pCMV $\beta$ gal. In both transfections, 2-4 x 10<sup>6</sup>  
25 HEK 293 cells in a 10-cm tissue culture plate were transiently co-transfected with 5  $\mu$ g of each of the plasmids included in the experiment according to standard CaPO<sub>4</sub> precipitation transfection procedures (Wigler et al.  
(1979) Proc. Natl. Acad. Sci. USA 76:1373-1376). The  
30 transfectants were analyzed for  $\beta$ -galactosidase expression by direct staining of the product of a reaction involving  $\beta$ -galactosidase and the X-gal substrate [Jones, J.R. (1986) EMBO 5:3133-3142] and by measurement of  $\beta$ -galactosidase activity [Miller, J.H. (1972) Experiments in Molecular  
35 Genetics, pp. 352-355, Cold Spring Harbor Press]. To

-78-

evaluate subunit cDNA expression in these transfectants, the cells were analyzed for subunit transcript producti n (northern analysis), subunit protein production (immunoblot analysis of cell lysates) and functional calcium chann 1  
5 expression (electrophysiological analysis).

**b. Stable transfection**

HEK 293 cells were transfected using the calcium phosphate transfection procedure [Current Protocols in Molecular Biology, Vol. 1, Wiley Inter-Science, Supplement 10 14, Unit 9.1.1-9.1.9 (1990)]. Ten-cm plates, each containing one-to-two million HEK 293 cells, were transfected with 1 ml of DNA/calcium phosphate precipitate containing 5 µg pVDCCIII(A), 5 µg pHBCaCH $\alpha_2$ A, 5µg pHBCaCH $\beta_1$ RBS(A), 5 µg pCMVBgal and 1 µg pSV2neo (as a 15 selectable marker). After 10-20 days of growth in media containing 500 µg G418, colonies had formed and were isolated using cloning cylinders.

20 2. Analysis of HEK 293 cells transiently transfected with DNA encoding human neuronal calcium channel subunits

**a. Analysis of  $\beta$ -galactosidase expression**

Transient transfectants were assayed for  $\beta$ -galactosidase expression by  $\beta$ -galactosidase activity assays (Miller, J.H., (1972) Experiments in Molecular Genetics, 25 pp. 352-355, Cold Spring Harbor Press) of cell lysates (prepared as described in Example VII.A.2) and staining of fixed cells (Jones, J.R. (1986) EMBO 5:3133-3142). The results of these assays indicated that approximately 30% of the HEK 293 cells had been transfected.

30 **b. Northern analysis**

PolyA+ RNA was isolated using the Invitrogen Fast Trak Kit (InVitrogen, San Diego, CA) from HEK 293 cells transiently transfected with DNA encoding each of the  $\alpha_1$ ,  $\alpha_2$  and  $\beta_1$  subunits and the lacZ gene or the  $\alpha_1$  subunit and the 35 lacZ gene. The RNA was subjected t electrophoresis on an agarose gel and transferred to nitr cellulose. The

-79-

nitrocellulose was then hybridized with one or more of the following radiolabeled probes: the lacZ gene, human neuronal calcium channel  $\alpha_1$  subunit-encoding cDNA, human neuronal calcium channel  $\alpha_2$  subunit-encoding cDNA or human neuronal calcium channel  $\beta_1$  subunit-encoding cDNA. Two transcripts that hybridized with the  $\alpha_1$  subunit-encoding cDNA were detected in HEK 293 cells transfected with the DNA encoding the  $\alpha_1$ ,  $\alpha_2$ , and  $\beta_1$  subunits and the lacZ gene as well as in HEK 293 cells transfected with the  $\alpha_1$  subunit cDNA and the lacZ gene. One mRNA species was the size expected for the transcript of the  $\alpha_1$  subunit cDNA (8000 nucleotides). The second RNA species was smaller (4000 nucleotides) than the size expected for this transcript. RNA of the size expected for the transcript of the lacZ gene was detected in cells transfected with the  $\alpha_1$ ,  $\alpha_2$  and  $\beta_1$  subunit-encoding cDNA and the lacZ gene and in cells transfected with the  $\alpha_1$  subunit cDNA and the lacZ gene by hybridization to the lacZ gene sequence.

RNA from cells transfected with the  $\alpha_1$ ,  $\alpha_2$  and  $\beta_1$  subunit-encoding cDNA and the lacZ gene was also hybridized with the  $\alpha_2$  and  $\beta_1$  subunit cDNA probes. Two mRNA species hybridized to the  $\alpha_2$  subunit cDNA probe. One species was the size expected for the transcript of the  $\alpha_2$  subunit cDNA (4000 nucleotides). The other species was larger (6000 nucleotides) than the expected size of this transcript. Multiple RNA species in the cells co-transfected with  $\alpha_1$ ,  $\alpha_2$  and  $\beta_1$  subunit-encoding cDNA and the lacZ gene hybridized to the  $\beta_1$  subunit cDNA probe. Multiple  $\beta$ -subunit transcripts of varying sizes were produced since the  $\beta$  subunit cDNA expression vector contains two potential polyA<sup>+</sup> addition sites.

#### c. Electrophysiological analysis

Individual transiently transfected HEK 293 cells were assayed for the presence of voltage-dependent barium currents using the whole-cell variant of the patch clamp

-80-

technique [Hamill et al. (1981). *Pflugers Arch.* 391:85-100]. HEK 293 cells transiently transfected with pCMV $\beta$ gal only were assayed for barium currents as a negative control in these experiments. The cells were placed in a bathing solution that contained barium ions to serve as the current carrier. Choline chloride, instead of NaCl or KCl, was used as the major salt component of the bath solution to eliminate currents through sodium and potassium channels. The bathing solution contained 1 mM MgCl<sub>2</sub>, and was buffered at pH 7.3 with 10 mM HEPES (pH adjusted with sodium or tetraethylammonium hydroxide). Patch pipettes were filled with a solution containing 135 mM CsCl, 1 mM MgCl<sub>2</sub>, 10 mM glucose, 10 mM EGTA, 4 mM ATP and 10 mM HEPES (pH adjusted to 7.3 with tetraethylammonium hydroxide). Cesium and tetraethylammonium ions block most types of potassium channels. Pipettes were coated with Sylgard (Dow-Corning, Midland, MI) and had resistances of 1-4 megohm. Currents were measured through a 500 megohm headstage resistor with the Axopatch IC (Axon Instruments, Foster City, CA) amplifier, interfaced with a Labmaster (Scientific Solutions, Solon, OH) data acquisition board in an IBM-compatible PC. PClamp (Axon Instruments) was used to generate voltage commands and acquire data. Data were analyzed with pClamp or Quattro Professional (Borland International, Scotts Valley, CA) programs.

To apply drugs, "puffer" pipettes positioned within several micrometers of the cell under study were used to apply solutions by pressure application. The drugs used for pharmacological characterization were dissolved in a solution identical to the bathing solution. Samples of a 10 mM stock solution of Bay K 8644 (RBI, Natick, MA), which was prepared in DMSO, were diluted to a final concentration of 1  $\mu$ M in 15 mM Ba<sup>2+</sup>-containing bath solution before they were applied.

-81-

Twenty-one negative control HEK 293 cells (transiently transfected with the lacZ gene expression vector pCMV $\beta$ gal only) were analyzed by the whole-cell variant of the patch clamp method for recording currents. Only one cell

5 displayed a discernable inward barium current; this current was not affected by the presence of 1  $\mu$ M Bay K 8644. In addition, application of Bay K 8644 to four cells that did not display Ba<sup>2+</sup> currents did not result in the appearance of any currents.

10 Two days after transient transfection of HEK 293 cells with  $\alpha_1$ ,  $\alpha_2$  and  $\beta_1$  subunit-encoding cDNA and the lacZ gene, individual transfectants were assayed for voltage-dependent barium currents. The currents in nine transfectants were recorded. Because the efficiency of transfection of one

15 cell can vary from the efficiency of transfection of another cell, the degree of expression of heterologous proteins in individual transfectants varies and some cells do not incorporate or express the foreign DNA. Inward barium currents were detected in two of these nine

20 transfectants. In these assays, the holding potential of the membrane was -90 mV. The membrane was depolarized in a series of voltage steps to different test potentials and the current in the presence and absence of 1  $\mu$ M Bay K 8644 was recorded. The inward barium current was significantly

25 enhanced in magnitude by the addition of Bay K 8644. The largest inward barium current (-160 pA) was recorded when the membrane was depolarized to 0 mV in the presence of 1  $\mu$ M Bay K 8644. A comparison of the I-V curves, generated by plotting the largest current recorded after each

30 depolarization versus the depolarization voltage, corresponding to recordings conducted in the absence and presence of Bay K 8644 illustrated the enhancement of the voltage-activated current in the presence of Bay K 8644.

Pronounced tail currents were detected in the tracings

35 of currents generated in the presence of Bay K 8644 in HEK

-82-

293 cells transfected with  $\alpha_1$ ,  $\alpha_2$  and  $\beta_1$  subunit-encoding cDNA and the lacZ gene, indicating that the recombinant calcium channels responsible for the voltage-activated barium currents recorded in this transfected appear to be  
5 DHP-sensitive.

The second of the two transfected cells that display d inward barium currents expressed a -50 pA current when the membrane was depolarized from -90 mV. This current was nearly completely blocked by 200  $\mu$ M cadmium, an established  
10 calcium channel blocker.  
10

Ten cells that were transiently transfected with the DNA encoding the  $\alpha_1$  subunit and the lacZ gene were analyzed by whole-cell patch clamp methods two days after transfection. One of these cells displayed a 30 pA inward  
15 barium current. This current amplified 2-fold in the presence of 1  $\mu$ M Bay K 8644. Furthermore, small tail currents were detected in the presence of Bay K 8644. These data indicate that expression of the human neuronal calcium channel  $\alpha_{1D}$  subunit-encoding cDNA in HEK 293 yields  
20 a functional DHP-sensitive calcium channel.  
20

3. Analysis of HEK 293 cells stably transfected with DNA encoding human neuronal calcium channel subunits

Individual stably transfected HEK 293 cells were  
25 assayed electrophysiologically for the presence of voltage-dependent barium currents as described for electrophysiological analysis of transiently transfected HEK 293 cells (see Example VII.B.2.c). In an effort to maximize calcium channel activity via cyclic-AMP-dependent  
30 kinase-mediated phosphorylation [Pelzer, et al. (1990) Rev. Physiol. Biochem. Pharmacol. 114:107-207], cAMP (Na salt,  
250  $\mu$ M) was added to the pipet solution and forskolin (10  $\mu$ M) was added to the bath solution in some of the recordings. Qualitatively similar results were obtained  
35 whether these compounds were present or not.  
35

-83-

Barium currents were recorded from stably transfected cells in the absence and presence of Bay K 8644 (1  $\mu$ M). When the cell was depolarized to -10 mV from a holding potential of -90 mV in the absence of Bay K 8644, a current of approximately 35 pA with a rapidly deactivating tail current was recorded. During application of Bay K 8644, an identical depolarizing protocol elicited a current of approximately 75 pA, accompanied by an augmented and prolonged tail current. The peak magnitude of currents recorded from this same cell as a function of a series of depolarizing voltages were assessed. The responses in the presence of Bay K 8644 not only increased, but the entire current-voltage relation shifted about -10 mV. Thus, three typical hallmarks of Bay K 8644 action, namely increased current magnitude, prolonged tail currents, and negatively shifted activation voltage, were observed, clearly indicating the expression of a DHP-sensitive calcium channel in these stably transfected cells. No such effects of Bay K 8644 were observed in untransfected HEK 293 cells, either with or without cAMP or forskolin.

C. Use of pCMV-based vectors and pcDNA1-based vectors for expression of DNA encoding human neuronal calcium channel subunits

1. Preparation of constructs

To determine if the levels of recombinant expression of human calcium channel subunit-encoding DNA in host cells could be enhanced by using pCMV-based instead of pcDNA1-based expression vectors, additional expression vectors were constructed. The full-length  $\alpha_1D$  cDNA from pVDCCIII(A) (see Example II.A.3.d), the full-length  $\alpha_1$  cDNA, contained on a 3600 bp EcoRI fragment from HBCaCH $\alpha_1$ , (see Example IV.B) and a full-length  $\beta_1$  subunit cDNA from pHBCaCH $\beta_1$ RBS(A) (see Example III.B.3) were separately subcloned into plasmid pCMV $\beta$ gal. Plasmid pCMV $\beta$ gal was digested with NotI to remove the lacZ gene. The remaining vector portion of the plasmid, referred to as pCMV, was blunt-ended at the NotI

-84-

sites. The full-length  $\alpha_2$ -encoding DNA and  $\beta_1$ -encoding DNA, contained on separate EcoRI fragments, were isolated, blunt-ended and separately ligated to the blunt-ended vector fragment of pCMV locating the cDNAs between the CMV promoter and SV40 polyadenylation sites in pCMV. To ligate the  $\alpha_{1D}$ -encoding cDNA with pCMV, the restriction sites in the polylinkers immediately 5' of the CMV promoter and immediately 3' of the SV40 polyadenylation site were removed from pCMV. A polylinker was added at the NotI site. The polylinker had the following sequence of restriction enzyme recognition sites:

15	GGCCGC	EcoRI	SalI	PstI	EcoRV	HindIII	XbaII	GT
	CG	site	site	site	site	site	site	CACCGG
	<i>NotI</i>							↑ Destroys Not

The  $\alpha_{1D}$ -encoding DNA, isolated as a BamHI/XhoI fragment from pVDCCIII(A), was then ligated to XbaII/SalI-digested pCMV to place it between the CMV promoter and SV40 polyadenylation site.

Plasmid pCMV contains the CMV promoter as does pcDNA1, but differs from pcDNA1 in the location of splice donor/splice acceptor sites relative to the inserted subunit-encoding DNA. After inserting the subunit-encoding DNA into pCMV, the splice donor/splice acceptor sites are located 3' of the CMV promoter and 5' of the subunit-encoding DNA start codon. After inserting the subunit-encoding DNA into pcDNA1, the splice donor/splice acceptor sites are located 3' of the subunit cDNA stop codon.

## 2. Transfection of HEK 293 cells

HEK 293 cells were transiently co-transfected with the  $\alpha_{1D}$ ,  $\alpha_2$  and  $\beta_1$  subunit-encoding DNA in pCMV or with the  $\alpha_{1D}$ ,  $\alpha_2$  and  $\beta$  subunit-encoding DNA in pcDNA1 (vectors pVDCCIII(A), pHBCaCH $\alpha_2$ A and pHBCaCH $\beta_1$ RBS(A), respectively), as described in Example VII.B.1.a. Plasmid pCMV $\beta$ gal was included in each transfection to as a measure of

-85-

transfection efficiency. The results of  $\beta$ -galactosidase assays of the transfectants (see Example VII.B.2.), indicated that HEK 293 cells were transfected equally efficiently with pCMV- and pcDNA1-based plasmids.

5           3. Northern analysis

Total and polyA<sup>+</sup> RNA were isolated from the transiently transfected cells as described in Examples VII.A.2 and VII.B.2.b. Northern blots of the RNA were hybridized with the following radiolabeled probes:  $\alpha_{1D}$  cDNA, human neuronal 10 calcium channel  $\alpha_1$  subunit cDNA and DNA encoding the human neuronal calcium channel  $\beta_1$  subunit. Messenger RNA of sizes expected for  $\alpha_{1D}$ ,  $\alpha_1$  and  $\beta_1$  subunit transcripts were detected in all transflectants. A greater amount of the  $\alpha_{1D}$  transcript was present in cells that were co-transfected 15 with pCMV-based plasmids than in cells that were co-transfected with pcDNA1-based plasmids. Equivalent amounts of  $\alpha_1$  and  $\beta_1$  subunit transcripts were detected in all transflectants.

20           D. Expression in *Xenopus laevis* oocytes of RNA encoding human neuronal calcium channel subunits

Various combinations of the transcripts of DNA encoding the human neuronal  $\alpha_{1D}$ ,  $\alpha_1$  and  $\beta_1$  subunits prepared in vitro were injected into *Xenopus laevis* oocytes. Those injected with combinations that included  $\alpha_{1D}$  exhibited 25 voltage-activated barium currents.

1. Preparation of transcripts

Transcripts encoding the human neuronal calcium channel  $\alpha_{1D}$ ,  $\alpha_1$  and  $\beta_1$  subunits were synthesized according to the instructions of the mCAP mRNA CAPPING KIT (Stratagene, 30 La Jolla, CA catalog #200350). Plasmids pVDCC III.RBS(A), containing of pcDNA1 and the  $\alpha_{1D}$  cDNA that begins with a ribosome binding site and the eighth ATG codon of the coding sequence (see Example III.A.3.d), plasmid pHBCaCH $\alpha_1$ A containing of pcDNA1 and an  $\alpha_1$  subunit cDNA (see Example 35 IV), and plasmid pHBCaCH $\beta_1$ RBS(A) containing pcDNA1 and the

-86-

$\beta_1$  DNA lacking intron sequence and containing a ribosome binding site (see Example III), were linearized by restriction digestion. The  $\alpha_1\text{D}$  cDNA- and  $\alpha_2$  subunit-encoding plasmids were digested with *Xba*I, and the  $\beta_1$  subunit-  
5 encoding plasmid was digested with *Eco*RV. The DNA insert was transcribed with T7 RNA polymerase.

## 2. Injection of oocytes

*Xenopus laevis* oocytes were isolated and defolliculated by collagenase treatment and maintained in  
10 100 mM NaCl, 2 mM KCl, 1.8 mM CaCl<sub>2</sub>, 1 mM MgCl<sub>2</sub>, 5 mM HEPES, pH 7.6, 20  $\mu$ g/ml ampicillin and 25  $\mu$ g/ml streptomycin at 19–25°C for 2 to 5 days after injection and prior to recording. For each transcript that was injected into the oocyte, 6 ng of the specific mRNA was injected per cell in  
15 a total volume of 50 nl.

## 3. Intracellular voltage recordings

Injected oocytes were examined for voltage-dependent barium currents using two-electrode voltage clamp methods [Dascal, N. (1987) CRC Crit. Rev. Biochem. 22:317]. The pClamp (Axon Instruments) software package was used in conjunction with a Labmaster 125 kHz data acquisition interface to generate voltage commands and to acquire and analyze data. Quattro Professional was also used in this analysis. Current signals were digitized at 1–5 kHz, and  
25 filtered appropriately. The bath solution contained of the following: 40 mM BaCl<sub>2</sub>, 36 mM tetraethylammonium chloride (TEA-Cl), 2 mM KCl, 5 mM 4-aminopyridine, 0.15 mM niflumic acid, 5 mM HEPES, pH 7.6.

### a. Electrophysiological analysis of oocytes injected with transcripts encoding the human neuronal calcium channel $\alpha_1$ , $\alpha_2$ and $\beta_1$ -subunits

Uninjected oocytes were examined by two-electrode voltage clamp methods and a very small (25 nA) endogenous  
35 inward Ba<sup>2+</sup> current was detected in only one of seven analyzed cells.

-87-

Oocytes coinjected with  $\alpha_{1D}$ ,  $\alpha_2$  and  $\beta_1$  subunit transcripts expressed sustained inward barium currents upon depolarization of the membrane from a holding potential of -90 mV or -50 mV ( $154 \pm 129$  nA, n=21). These currents typically showed little inactivation when test pulses ranging from 140 to 700 msec. were administered. Depolarization to a series of voltages revealed currents that first appeared at approximately -30 mV and peaked at approximately 0 mV.

Application of the DHP Bay K 8644 increased the magnitude of the currents, prolonged the tail currents present upon repolarization of the cell and induced a hyperpolarizing shift in current activation. Bay K 8644 was prepared fresh from a stock solution in DMSO and introduced as a 10x concentrate directly into the 60  $\mu$ l bath while the perfusion pump was turned off. The DMSO concentration of the final diluted drug solutions in contact with the cell never exceeded 0.1%. Control experiments showed that 0.1% DMSO had no effect on membrane currents.

Application of the DHP antagonist nifedipine (stock solution prepared in DMSO and applied to the cell as described for application of Bay K 8644) blocked a substantial fraction ( $91 \pm 6\%$ , n=7) of the inward barium current in oocytes coinjected with transcripts of the  $\alpha_{1D}$ ,  $\alpha_2$  and  $\beta_1$  subunits. A residual inactivating component of the inward barium current typically remained after nifedipine application. The inward barium current was blocked completely by 50  $\mu$ M Cd<sup>2+</sup>, but only approximately 15% by 100  $\mu$ M Ni<sup>2+</sup>.

The effect of  $\omega$ CgTX on the inward barium currents in oocytes co-injected with transcripts of the  $\alpha_{1D}$ ,  $\alpha_2$ , and  $\beta_1$  subunits was investigated.  $\omega$ CgTX (Bachem, Inc., Torrance CA) was prepared in the 15 mM BaCl<sub>2</sub> bath solution plus 0.1% cytochrome C (Sigma) to serve as a carrier protein.

-88-

Control experiments showed that cytochrome C had no effect on currents. A series of voltage pulses from a -90 mV holding potential to 0 mV were recorded at 20 msec. intervals. To reduce the inhibition of  $\omega$ CgTX binding by 5 divalent cations, recordings were made in 15 mM BaCl<sub>2</sub>, 73.5 mM tetraethylammonium chloride, and the remaining ingredients identical to the 40 mM Ba<sup>2+</sup> recording solution. Bay K 8644 was applied to the cell prior to addition to  $\omega$ CgTX in order to determine the effect of  $\omega$ CgTX on the DHP- 10 sensitive current component that was distinguished by the prolonged tail currents. The inward barium current was blocked weakly ( $54 \pm 29\%$ , n=7) and reversibly by relatively high concentrations (10-15  $\mu$ M) of  $\omega$ CgTX. The test currents and the accompanying tail currents were blocked 15 progressively within two to three minutes after application of  $\omega$ CgTX, but both recovered partially as the  $\omega$ CgTX was flushed from the bath.

20 b. Analysis of oocytes injected with only a transcripts encoding the human neuronal calcium channel  $\alpha_{1D}$  or transcripts encoding an  $\alpha_{1D}$  and other subunits

The contribution of the  $\alpha_2$  and  $\beta_1$  subunits to the inward barium current in oocytes injected with transcripts 25 encoding the  $\alpha_{1D}$ ,  $\alpha_2$  and  $\beta_1$  subunits was assessed by expression of the  $\alpha_{1D}$  subunit alone or in combination with either the  $\beta_1$  subunit or the  $\alpha_2$  subunit. In oocytes injected with only the transcript of a  $\alpha_{1D}$  cDNA, no Ba<sup>2+</sup> currents were detected (n=3). In oocytes injected with 30 transcripts of  $\alpha_{1D}$  and  $\beta_1$  cDNAs, small ( $108 \pm 39$  nA) Ba<sup>2+</sup> currents were detected upon depolarization of the membrane from a holding potential of -90 mV that resembled the currents observed in cells injected with transcripts of  $\alpha_{1D}$ ,  $\alpha_2$  and  $\beta_1$  cDNAs, although the magnitude of the current was 35 less. In two of the four oocytes injected with transcripts of the  $\alpha_{1D}$ -encoding and  $\beta_1$ -encoding DNA, the Ba<sup>2+</sup> currents

-89-

exhibited a sensitivity to Bay K 8644 that was similar to the Bay K 8644 sensitivity of Ba<sup>2+</sup> currents expressed in oocytes injected with transcripts encoding the  $\alpha_{1D}$   $\alpha_1$ -,  $\alpha_2$  and  $\beta_1$  subunits.

5 Three of five oocytes injected with transcripts encoding the  $\alpha_{1D}$  and  $\alpha_2$  subunits exhibited very small Ba<sup>2+</sup> currents (15-30 nA) upon depolarization of the membrane from a holding potential of -90 mV. These barium currents showed little or no response to Bay K 8644.

10 c. Analysis of oocytes injected with transcripts encoding the human neuronal calcium channel  $\alpha_2$  and/or  $\beta_1$  subunit

To evaluate the contribution of the  $\alpha_{1D}$   $\alpha_1$ -subunit to the inward barium currents detected in oocytes co-injected 15 with transcripts encoding the  $\alpha_{1D}$ ,  $\alpha_2$  and  $\beta_1$  subunits, oocytes injected with transcripts encoding the human neuronal calcium channel  $\alpha_2$  and/or  $\beta_1$  subunits were assayed for barium currents. Oocytes injected with transcripts encoding the  $\alpha_2$  subunit displayed no detectable inward 20 barium currents (n=5). Oocytes injected with transcripts encoding a  $\beta_1$  subunit displayed measurable (54 ± 23 nA, n=5) inward barium currents upon depolarization and oocytes injected with transcripts encoding the  $\alpha_2$  and  $\beta_1$  subunits displayed inward barium currents that were approximately 25 50% larger (80 ± 61 nA, n=18) than those detected in oocytes injected with transcripts of the  $\beta_1$ -encoding DNA only.

The inward barium currents in oocytes injected with transcripts encoding the  $\beta_1$  subunit or  $\alpha_2$  and  $\beta_1$  subunits 30 typically were first observed when the membrane was depolarized to -30 mV from a holding potential of -90 mV and peaked when the membrane was depolarized to 10 to 20 mV. Macroscopically, the currents in oocytes injected with transcripts encoding the  $\alpha_2$  and  $\beta_1$  subunits or with 35 transcripts encoding the  $\beta_1$  subunit were indistinguishable. In contrast to the currents in oocytes co-injected with

-90-

transcripts of  $\alpha_{1D}$ ,  $\alpha_2$  and  $\beta_1$  subunit cDNAs, these currents showed a significant inactivation during the test pulse and a strong sensitivity to the holding potential. The inward barium currents in oocytes co-injected with transcripts encoding the  $\alpha_2$  and  $\beta_1$  subunits usually inactivated to 10-60% of the peak magnitude during a 140-msec pulse and were significantly more sensitive to holding potential than those in oocytes co-injected with transcripts encoding the  $\alpha_{1D}$ ,  $\alpha_2$  and  $\beta_1$  subunits. Changing the holding potential of the membranes of oocytes co-injected with transcripts encoding the  $\alpha_2$  and  $\beta_1$  subunits from -90 to -50 mV resulted in an approximately 81% (n=11) reduction in the magnitude of the inward barium current of these cells. In contrast, the inward barium current measured in oocytes co-injected with transcripts encoding the  $\alpha_{1D}$ ,  $\alpha_2$  and  $\beta_1$  subunits were reduced approximately 24% (n=11) when the holding potential was changed from -90 to -50 mV.

The inward barium currents detected in oocytes injected with transcripts encoding the  $\alpha_2$  and  $\beta_1$  subunits were pharmacologically distinct from those observed in oocytes co-injected with transcripts encoding the  $\alpha_{1D}$ ,  $\alpha_2$  and  $\beta_1$  subunits. Oocytes injected with transcripts encoding the  $\alpha_2$  and  $\beta_1$  subunits displayed inward barium currents that were insensitive to Bay K 8644 (n=11). Nifedipine sensitivity was difficult to measure because of the holding potential sensitivity of nifedipine and the current observed in oocytes injected with transcripts encoding the  $\alpha_2$  and  $\beta_1$  subunits. Nevertheless, two oocytes that were co-injected with transcripts encoding the  $\alpha_2$  and  $\beta_1$  subunits displayed measurable (25 to 45 nA) inward barium currents when depolarized from a holding potential of -50 mV. These currents were insensitive to nifedipine (5 to 10  $\mu$ M). The inward barium currents in oocytes injected with transcripts encoding the  $\alpha_2$  and  $\beta_1$  subunits showed the same sensitivity to heavy metals as the currents detected in oocytes

-91-

injected with transcripts encoding the  $\alpha_{1D}$ ,  $\alpha_2$  and  $\beta_1$  subunits.

The inward barium current detected in oocytes injected with transcripts encoding the human neuronal  $\alpha_2$  and  $\beta_1$  subunits has pharmacological and biophysical properties that resemble calcium currents in uninjected *Xenopus* oocytes. Because the amino acids of this human neuronal calcium channel  $\beta_1$  subunit lack hydrophobic segments capable of forming transmembrane domains, it is unlikely that recombinant  $\beta_1$  subunits alone can form an ion channel. It is more probable that a homologous endogenous  $\alpha_1$  subunit exists in oocytes and that the activity mediated by such an  $\alpha_1$  subunit is enhanced by expression of a human neuronal  $\beta_1$  subunit.

While the invention has been described with some specificity, modifications apparent to those with ordinary skill in the art may be made without departing from the scope of the invention. Since such modifications will be apparent to those of skill in the art, it is intended that this invention be limited only by the scope of the appended claims.

-92-

SEQUENCE LISTING

(1) GENERAL INFORMATION:

(i) APPLICANT: Harpold, Michael  
Ellis, Steven  
Williams, Mark  
Feldman, Daniel  
McCue, Ann  
Brenner, Robert

(ii) TITLE OF INVENTION: HUMAN CALCIUM CHANNEL COMPOSITIONS AND METHODS

(iii) NUMBER OF SEQUENCES: 21

(iv) CORRESPONDENCE ADDRESS:

(A) ADDRESSEE: Fitch, Even, Tabin & Flannery  
(B) STREET: 4250 Executive Square, Suite 510  
(C) CITY: La Jolla  
(D) STATE: CA  
(F) ZIP: 92037

(v) COMPUTER READABLE FORM:

(A) MEDIUM TYPE: Floppy disk  
(B) COMPUTER: IBM PC compatible  
(C) OPERATING SYSTEM: PC-DOS/MS-DOS  
(D) SOFTWARE: PatentIn Release #1.0, Version #1.25

(vi) CURRENT APPLICATION DATA:

(A) APPLICATION NUMBER:  
(B) FILING DATE:  
(C) CLASSIFICATION:

(vii) PRIOR APPLICATION DATA:

(A) APPLICATION NUMBER: US 07/868,354  
(B) FILING DATE: 10-APR-1992

(vii) PRIOR APPLICATION DATA:

(A) APPLICATION NUMBER: US 07/745,206  
(B) FILING DATE: 15-AUG-1991

(vii) PRIOR APPLICATION DATA:

(A) APPLICATION NUMBER: US 07/620,250  
(B) FILING DATE: 30-NOV-1990

(vii) PRIOR APPLICATION DATA:

(A) APPLICATION NUMBER: US 07/482,384  
(B) FILING DATE: 20-FEB-1990

(vii) PRIOR APPLICATION DATA:

(A) APPLICATION NUMBER: US 07/603,751  
(B) FILING DATE: 04-APR-1989

(vii) PRIOR APPLICATION DATA:

(A) APPLICATION NUMBER: WO PCT/US89/01408  
(B) FILING DATE: 04-APR-1989

(vii) PRIOR APPLICATION DATA:

(A) APPLICATION NUMBER: US 07/176,899  
(B) FILING DATE: 04-APR-1988

(viii) ATTORNEY/AGENT INFORMATION:

(A) NAME: Seidman, Stephanie L.  
(B) REGISTRATION NUMBER: 33,779  
(C) REFERENCE/DOCKET NUMBER: 53607PCT

-93-

## (ix) TELECOMMUNICATION INFORMATION:

- (A) TELEPHONE: 619-552-1311
- (B) TELEFAX: 619-552-0095

## (2) INFORMATION FOR SEQ ID NO:1:

## (i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 7635 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: double
- (D) TOPOLOGY: linear

## (ii) MOLECULE TYPE: DNA (genomic)

## (ix) FEATURE:

- (A) NAME/KEY: CDS
- (B) LOCATION: 511..6996

## (ix) FEATURE:

- (A) NAME/KEY: 5'UTR
- (B) LOCATION: 1..510

## (ix) FEATURE:

- (A) NAME/KEY: 3'UTR
- (B) LOCATION: 6994..7635

## (xi) SEQUENCE DESCRIPTION: SEQ ID NO:1:

GGCGAGCGC	CTCCGTCCCC	GGATGTGAGC	TCCGGCTGCC	CGCGGTCCCG	AGCCAGCGGC	60
GCGCGGGCGG	CGCGGGCGGG	CACCGGGCAC	CGCGGCCGGC	GGGCAGACGG	GCGGGCATGG	120
GGGGAGCGCC	GAGCGGCCCC	GGCGGCCGGG	CCGGCATCAC	CGCGCGTCT	CTCCGCTAGA	180
GGAGGGGACA	AGCCAGTTCT	CCTTTGCAGC	AAAAAATTAC	ATGTATATAT	TATTAAGATA	240
ATATATACAT	TGGATTTTAT	TTTTTTAAAA	AGTTTATTT	GCTCCATTTC	TGAAAAAGAG	300
AGAGCTTGGG	TGGCGAGCGG	TTTTTTTTA	AAATCAATT	TCCTTATTT	CTGTTATTTG	360
TCCCCGTCCC	TCCCCACCCC	CCTGCTGAAG	CGAGAATAAG	GGCAGGGACC	GCGGCTCCTA	420
CCTCTTGGTG	ATCCCCTTCC	CCATTCCGCC	CCCCCCCCAA	GGCCCAGCAC	AGTGCCTG	480
ACACAGTAGT	CGCTCAATAA	ATGTTCGTGG	ATG ATG ATG ATG ATG ATG AAA			534
			Met Met Met Met Met Met Lys	1	5	
AAA ATG CAG CAT CAA CGG CAG CAG CAA GCG GAC CAC GCG AAC GAG GCA						
Lys Met Gln His Gln Arg Gln Gln Gln Ala Asp His Ala Asn Glu Ala	10	15	20			582
AAC TAT GCA AGA GGC ACC AGA CTT CCT CTT TCT GGT GAA GGA CCA ACT						
Asn Tyr Ala Arg Gly Thr Arg Leu Pro Leu Ser Gly Glu Gly Pro Thr	25	30	35	40		630
TCT CAG CCG AAT AGC TCC AAG CAA ACT GTC CTG TCT TGG CAA GCT GCA						
Ser Gln Pro Asn Ser Ser Lys Gln Thr Val Leu Ser Trp Gln Ala Ala	45	50	55			678
ATC GAT GCT GCT AGA CAG GCC AAG GCT GCC CAA ACT ATG AGC ACC TCT						
II Asp Ala Ala Arg Gln Ala Lys Ala Ala Gln Thr Met Ser Thr Ser	60	65	70			726
GCA CCC CCA CCT GTA GGA TCT CTC TCC CAA AGA AAA CGT CAG CAA TAC						
Ala Pro Pro Pro Val Gly Ser Leu Ser Gln Arg Lys Arg Gln Gln Tyr	75	80	85			774

-94-

GCC AAG AGC AAA AAA CAG GGT AAC TCG TCC AAC AGC CGA CCT GCC CGC Ala Lys Ser Lys Lys Gln Gly Asn Ser Ser Asn Ser Arg Pro Ala Arg 90 95 100	822
GCC CTT TTC TGT TTA TCA CTC AAT AAC CCC ATC CGA AGA GCC TGC ATT Ala Leu Phe Cys Leu Ser Leu Asn Asn Pro Ile Arg Arg Ala Cys Ile 105 110 115 120	870
AGT ATA GTG GAA TGG AAA CCA TTT GAC ATA TTT ATA TTA TTG GCT ATT Ser Ile Val Glu Trp Lys Pro Phe Asp Ile Phe Ile Leu Leu Ala Ile 125 130 135	918
TTT GCC AAT TGT GTG GCC TTA GCT ATT TAC ATC CCA TTC CCT GAA GAT Phe Ala Asn Cys Val Ala Leu Ala Ile Tyr Ile Pro Phe Pro Glu Asp 140 145 150	966
GAT TCT AAT TCA ACA AAT CAT AAC TTG GAA AAA GTA GAA TAT GCC TTC Asp Ser Asn Ser Thr Asn His Asn Leu Glu Lys Val Glu Tyr Ala Phe 155 160 165	1014
CTG ATT ATT TTT ACA GTC GAG ACA TTT TTG AAG ATT ATA GCG TAT GGA Leu Ile Ile Phe Thr Val Glu Thr Phe Leu Lys Ile Ile Ala Tyr Gly 170 175 180	1062
TTA TTG CTA CAT CCT AAT GCT TAT GTT AGG AAT GGA TGG AAT TTA CTG Leu Leu Leu His Pro Asn Ala Tyr Val Arg Asn Gly Trp Asn Leu Leu 185 190 195 200	1110
GAT TTT GTT ATA GTA ATA GTA GGA TTG TTT AGT GTA ATT TTG GAA CAA Asp Phe Val Ile Val Ile Val Gly Leu Phe Ser Val Ile Leu Glu Gln 205 210 215	1158
TTA ACC AAA GAA ACA GAA GGC GGG AAC CAC TCA AGC GGC AAA TCT GGA Leu Thr Lys Glu Thr Glu Gly Asn His Ser Ser Gly Lys Ser Gly 220 225 230	1206
GGC TTT GAT GTC AAA GCC CTC CGT GCC TTT CGA GTG TTG CGA CCA CTT Gly Phe Asp Val Lys Ala Leu Arg Ala Phe Arg Val Leu Arg Pro Leu 235 240 245	1254
CGA CTA GTG TCA GGA GTG CCC AGT TTA CAA GTT GTC CTG AAC TCC ATT Arg Leu Val Ser Gly Val Pro Ser Leu Gln Val Leu Asn Ser Ile 250 255 260	1302
ATA AAA GCC ATG GTT CCC CTC CTT CAC ATA GCC CTT TTG GTA TTA TTT Ile Lys Ala Met Val Pro Leu Leu His Ile Ala Leu Leu Val Leu Phe 265 270 275 280	1350
GTA ATC ATA ATC TAT GCT ATT ATA GGA TTG GAA CTT TTT ATT GGA AAA Val Ile Ile Ile Tyr Ala Ile Ile Gly Leu Glu Leu Phe Ile Gly Lys 285 290 295	1398
ATG CAC AAA ACA TGT TTT TTT GCT GAC TCA GAT ATC GTA GCT GAA GAG Met His Lys Thr Cys Phe Phe Ala Asp Ser Asp Ile Val Ala Glu Glu 300 305 310	1446
GAC CCA GCT CCA TGT GCG TTC TCA GGG AAT GGA CGC CAG TGT ACT GCC Asp Pro Ala Pro Cys Ala Phe Ser Gly Asn Gly Arg Gln Cys Thr Ala 315 320 325	1494
AAT GGC ACG GAA TGT AGG AGT GGC TGG GTT GGC CCG AAC GGA GGC ATC Asn Gly Thr Glu Cys Arg Ser Gly Trp Val Gly Pro Asn Gly Gly Ile 330 335 340	1542
ACC AAC TTT GAT AAC TTT GCC TTT GCC ATG CTT ACT GTG TTT CAG TGC Thr Asn Phe Asp Asn Phe Ala Phe Ala Met Leu Thr Val Phe Gln Cys 345 350 355 360	1590

-95-

ATC ACC ATG GAG GGC TGG ACA GAC GTG CTC TAC TGG ATG AAT GAT GCT Ile Thr Met Glu Gly Trp Thr Asp Val Leu Tyr Trp Met Asn Asp Ala 365 370 375	1638
ATG GGA TTT GAA TTG CCC TGG GTG TAT TTT GTC AGT CTC GTC ATC TTT Met Gly Phe Glu Leu Pro Trp Val Tyr Phe Val Ser Leu Val Ile Phe 380 385 390	1686
GGG TCA TTT TTC GTA CTA AAT CTT GTA CTT GGT GTA TTG AGC GGA GAA Gly Ser Phe Phe Val Leu Asn Leu Val Leu Gly Val Leu Ser Gly Glu 395 400 405	1734
TTC TCA AAG GAA AGA GAG AAG GCA AAA GCA CGG GGA GAT TTC CAG AAG Phe Ser Lys Glu Arg Glu Lys Ala Lys Ala Arg Gly Asp Phe Gln Lys 410 415 420	1782
CTC CGG GAG AAG CAG CAG CTG GAG GAT CTA AAG GGC TAC TTG GAT Leu Arg Glu Lys Gln Leu Glu Glu Asp Leu Lys Gly Tyr Leu Asp 425 430 435 440	1830
TGG ATC ACC CAA GCT GAG GAC ATC GAT CCG GAG AAT GAG GAA GAA GGA Trp Ile Thr Gln Ala Glu Asp Ile Asp Pro Glu Asn Glu Glu Gly 445 450 455	1878
GGA GAG GAA GGC AAA CGA AAT ACT AGC ATG CCC ACC AGC GAG ACT GAG Gly Glu Gly Lys Arg Asn Thr Ser Met Pro Thr Ser Glu Thr Glu 460 465 470	1926
TCT GTG AAC ACA GAG AAC GTC AGC GGT GAA GGC GAG AAC CGA GGC TGC Ser Val Asn Thr Glu Asn Val Ser Gly Glu Gly Glu Asn Arg Gly Cys 475 480 485	1974
TGT GGA AGT CTC TGT CAA GCC ATC TCA AAA TCC AAA CTC AGC CGA CGC Cys Gly Ser Leu Cys Gln Ala Ile Ser Lys Ser Lys Leu Ser Arg Arg 490 495 500	2022
TGG CGT CGC TGG AAC CGA TTC AAT CGC AGA AGA TGT AGG GCC GCC GTG Trp Arg Arg Trp Asn Arg Phe Asn Arg Arg Cys Arg Ala Ala Val 505 510 515 520	2070
AAG TCT GTC ACG TTT TAC TGG CTG GTT ATC GTC CTG GTG TTT CTG AAC Lys Ser Val Thr Phe Tyr Trp Leu Val Ile Val Leu Val Phe Leu Asn 525 530 535	2118
ACC TTA ACC ATT TCC TCT GAG CAC TAC AAT CAG CCA GAT TGG TTG ACA Thr Leu Thr Ile Ser Ser Glu His Tyr Asn Gln Pro Asp Trp Leu Thr 540 545 550	2166
CAG ATT CAA GAT ATT GCC AAC AAA GTC CTC TTG GCT CTG TTC ACC TGC Gln Ile Gln Asp Ile Ala Asn Lys Val Leu Leu Ala Leu Phe Thr Cys 555 560 565	2214
GAG ATG CTG GTA AAA ATG TAC AGC TTG GGC CTC CAA GCA TAT TTC GTC Glu Met Leu Val Lys Met Tyr Ser Leu Gly Leu Gln Ala Tyr Phe Val 570 575 580	2262
TCT CTT TTC AAC CGG TTT GAT TGC TTC GTG GTG TGT GGT GGA ATC ACT Ser Leu Phe Asn Arg Phe Asp Cys Phe Val Val Cys Gly Gly Ile Thr 585 590 595 600	2310
GAG ACG ATC TTG GTG GAA CTG GAA ATC ATG TCT CCC CTG GGG ATC TCT Glu Thr Ile Leu Val Glu Leu Glu Ile Met Ser Pro Leu Gly Ile Ser 605 610 615	2358
GTG TTT CGG TGT GTG CGC CTC TTA AGA ATC TTC AAA GTG ACC AGG CAC Val Phe Arg Cys Val Arg Leu Leu Arg Ile Phe Lys Val Thr Arg His 620 625 630	2406

-96-

TGG ACT TCC CTG AGC AAC TTA GTG GCA TCC TTA TTA AAC TCC ATG AAG Trp Thr Ser Leu Ser Asn Leu Val Ala Ser Leu Leu Asn Ser Met Lys 635 640 645	2454
TCC ATC GCT TCG CTG TTG CTT CTG CTT TTT CTC TTC ATT ATC ATC TTT Ser Ile Ala Ser Leu Leu Leu Leu Phe Leu Phe Ile Ile Ile Phe 650 655 660	2502
TCC TTG CTT GGG ATG CAG CTG TTT GGC GGC AAG TTT AAT TTT GAT GAA Ser Leu Leu Gly Met Gln Leu Phe Gly Gly Lys Phe Asn Phe Asp Glu 665 670 675 680	2550
ACG CAA ACC AAG CGG AGC ACC TTT GAC AAT TTC CCT CAA GCA CTT CTC Thr Gln Thr Lys Arg Ser Thr Phe Asp Asn Phe Pro Gln Ala Leu Leu 685 690 695	2598
ACA GTG TTC CAG ATC CTG ACA GGC GAA GAC TGG AAT GCT GTG ATG TAC Thr Val Phe Gln Ile Leu Thr Gly Glu Asp Trp Asn Ala Val Met Tyr 700 705 710	2646
GAT GGC ATC ATG GCT TAC GGG GGC CCA TCC TCT TCA GGA ATG ATC GTC Asp Gly Ile Met Ala Tyr Gly Pro Ser Ser Ser Gly Met Ile Val 715 720 725	2694
TGC ATC TAC TTC ATC ATC CTC TTC ATT TGT GGT AAC TAT ATT CTA CTG Cys Ile Tyr Phe Ile Ile Leu Phe Ile Cys Gly Asn Tyr Ile Leu Leu 730 735 740	2742
AAT GTC TTC TTG GCC ATC GCT GTA GAC AAT TTG GCT GAT GCT GAA AGT Asn Val Phe Leu Ala Ile Ala Val Asp Asn Leu Ala Asp Ala Glu Ser 745 750 755 760	2790
CTG AAC ACT GCT CAG AAA GAA GAA GCG GAA GAA AAG GAG AGG AAA AAG Leu Asn Thr Ala Gln Lys Glu Glu Ala Glu Glu Lys Glu Arg Lys Lys 765 770 775	2838
ATT GCC AGA AAA GAG AGC CTA GAA AAT AAA AAG AAC AAC AAA CCA GAA Ile Ala Arg Lys Glu Ser Leu Glu Asn Lys Asn Asn Lys Pro Glu 780 785 790	2886
GTC AAC CAG ATA GCC AAC AGT GAC AAC AAG GTT ACA ATT GAT GAC TAT Val Asn Gln Ile Ala Asn Ser Asp Asn Lys Val Thr Ile Asp Asp Tyr 795 800 805	2934
AGA GAA GAG GAT GAA GAC AAG GAC CCC TAT CCG CCT TGC GAT GTG CCA Arg Glu Glu Asp Glu Asp Lys Asp Pro Tyr Pro Pro Cys Asp Val Pro 810 815 820	2982
GTA GGG GAA GAG GAA GAG GAA GAG GAG GAG GAT GAA CCT GAG GTT CCT Val Gly Glu Glu Glu Glu Glu Glu Asp Glu Pro Glu Val Pro 825 830 835 840	3030
GCC GGA CCC CGT CCT CGA AGG ATC TCG GAG TTG AAC ATG AAG GAA AAA Ala Gly Pro Arg Pro Arg Arg Ile Ser Glu Leu Asn Met Lys Glu Lys 845 850 855	3078
ATT GCC CCC ATC CCT GAA GGG AGC GCT TTC TTC ATT CTT AGC AAG ACC Ile Ala Pro Ile Pro Glu Gly Ser Ala Phe Phe Ile Leu Ser Lys Thr 860 865 870	3126
AAC CCG ATC CGC GTA GGC TGC CAC AAG CTC ATC AAC CAC CAC ATC TTC Asn Pro Ile Arg Val Gly Cys His Lys Leu Ile Asn His His Ile Phe 875 880 885	3174
ACC AAC CTC ATC CTT GTC TTC ATC ATG CTG AGC AGT GCT GCC CTG GCC Thr Asn Leu Ile Leu Val Phe Ile Met Leu Ser Ser Ala Ala Leu Ala 890 895 900	3222

-97-

GCA GAG GAC CCC ATC CGC AGC CAC TCC TTC CGG AAC ACG ATA CTG GGT Ala Glu Asp Pro Ile Arg Ser His Ser Phe Arg Asn Thr Ile Leu Gly 905 910 915 920	3270
TAC TTT GAC TAT GCC TTC ACA GCC ATC TTT ACT GTT GAG ATC CTG TTG Tyr Phe Asp Tyr Ala Phe Thr Ala Ile Phe Thr Val Glu Ile Leu Leu 925 930 935	3318
AAG ATG ACA ACT TTT GGA GCT TTC CTC CAC AAA GGG GCC TTC TGC AGG Lys Met Thr Thr Phe Gly Ala Phe Leu His Lys Gly Ala Phe Cys Arg 940 945 950	3366
AAC TAC TTC AAT TTG CTG GAT ATG CTG GTG GTT GGG GTG TCT CTG GTG Asn Tyr Phe Asn Leu Leu Asp Met Leu Val Val Gly Val Ser Leu Val 955 960 965	3414
TCA TTT GGG ATT CAA TCC AGT GCC ATC TCC GTT GTG AAG ATT CTG AGG Ser Phe Gly Ile Gln Ser Ser Ala Ile Ser Val Val Lys Ile Leu Arg 970 975 980	3462
GTC TTA AGG GTC CTG CGT CCC CTC AGG GCC ATC AAC AGA GCA AAA GGA Val Leu Arg Val Leu Arg Pro Leu Arg Ala Ile Asn Arg Ala Lys Gly 985 990 995 1000	3510
CTT AAG CAC GTG GTC CAG TGC GTC TTC GTG GCC ATC CGG ACC ATC GGC Leu Lys His Val Val Gln Cys Val Phe Val Ala Ile Arg Thr Ile Gly 1005 1010 1015	3558
AAC ATC ATG ATC GTC ACC ACC CTC CTG CAG TTC ATG TTT GCC TGT ATC Asn Ile Met Ile Val Thr Thr Leu Leu Gln Phe Met Phe Ala Cys Ile 1020 1025 1030	3606
GGG GTC CAG TTG TTC AAG GGG AAG TTC TAT CGC TGT ACG GAT GAA GCC Gly Val Gln Leu Phe Lys Gly Lys Phe Tyr Arg Cys Thr Asp Glu Ala 1035 1040 1045	3654
AAA AGT AAC CCT GAA GAA TGC AGG GGA CTT TTC ATC CTC TAC AAG GAT Lys Ser Asn Pro Glu Glu Cys Arg Gly Leu Phe Ile Leu Tyr Lys Asp 1050 1055 1060	3702
GGG GAT GTT GAC AGT CCT GTG GTC CGT GAA CGG ATC TGG CAA AAC AGT Gly Asp Val Asp Ser Pro Val Val Arg Glu Arg Ile Trp Gln Asn Ser 1065 1070 1075 1080	3750
GAT TTC AAC TTC GAC AAC GTC CTC TCT GCT ATG ATG GCG CTC TTC ACA Asp Phe Asn Phe Asp Asn Val Leu Ser Ala Met Met Ala Leu Phe Thr 1085 1090 1095	3798
GTC TCC ACG TTT GAG GGC TGG CCT GCG TTG CTG TAT AAA GCC ATC GAC Val Ser Thr Phe Glu Gly Trp Pro Ala Leu Leu Tyr Lys Ala Ile Asp 1100 1105 1110	3846
TCG AAT GGA GAG AAC ATC GGC CCA ATC TAC AAC CAC CGC GTG GAG ATC Ser Asn Gly Glu Asn Ile Gly Pro Ile Tyr Asn His Arg Val Glu Ile 1115 1120 1125	3894
TCC ATC TTC TTC ATC ATC TAC ATC ATC ATT GTA GCT TTC TTC ATG ATG Ser Ile Phe Phe Ile Ile Tyr Ile Ile Ile Val Ala Phe Phe Met Met 1130 1135 1140	3942
AAC ATC TTT GTG GGC TTT GTC ATC GTT ACA TTT CAG GAA CAA GGA GAA Asn Ile Phe Val Gly Phe Val Ile Val Thr Phe Gln Glu Gln Gly Glu 1145 1150 1155 1160	3990
AAA GAG TAT AAG AAC TGT GAG CTG GAC AAA AAT CAG CGT CAG TGT GTT Lys Glu Tyr Lys Asn Cys Glu Leu Asp Lys Asn Gln Arg Gln Cys Val 1165 1170 1175	4038

-98-

GAA TAC GCC TTG AAA GCA CGT CCC TTG CCG AGA TAC ATC CCC AAA AAC Glu Tyr Ala Leu Lys Ala Arg Pro Leu Arg Arg Tyr Ile Pro Lys Asn 1180 1185 1190	4086
CCC TAC CAG TAC AAG TTC TGG TAC GTG GTG AAC TCT TCG CCT TTC GAA Pro Tyr Gln Tyr Lys Phe Trp Tyr Val Val Asn Ser Ser Pro Phe Glu 1195 1200 1205	4134
TAC ATG ATG TTT GTC CTC ATC ATG CTC AAC ACA CTC TGC TTG GCC ATG Tyr Met Met Phe Val Leu Ile Met Leu Asn Thr Leu Cys Leu Ala Met 1210 1215 1220	4182
CAG CAC TAC GAG CAG TCC AAG ATG TTC AAT GAT GCC ATG GAC ATT CTG Gln His Tyr Glu Gln Ser Lys Met Phe Asn Asp Ala Met Asp Ile Leu 1225 1230 1235 1240	4230
AAC ATG GTC TTC ACC GGG GTG TTC ACC GTC GAG ATG GTT TTG AAA GTC Asn Met Val Phe Thr Gly Val Phe Thr Val Glu Met Val Leu Lys Val 1245 1250 1255	4278
ATC GCA TTT AAG CCT AAG GGG TAT TTT AGT GAC GCC TGG AAC ACG TTT Ile Ala Phe Lys Pro Lys Gly Tyr Phe Ser Asp Ala Trp Asn Thr Phe 1260 1265 1270	4326
GAC TCC CTC ATC GTA ATC GGC AGC ATT ATA GAC GTG GCC CTC AGC GAA Asp Ser Leu Ile Val Ile Gly Ser Ile Ile Asp Val Ala Leu Ser Glu 1275 1280 1285	4374
GCA GAC CCA ACT GAA AGT GAA AAT GTC CCT GTC CCA ACT GCT ACA CCT Ala Asp Pro Thr Glu Ser Glu Asn Val Pro Val Pro Thr Ala Thr Pro 1290 1295 1300	4422
GGG AAC TCT GAA GAG AGC AAT AGA ATC TCC ATC ACC TTT TTC CGT CTT Gly Asn Ser Glu Glu Ser Asn Arg Ile Ser Ile Thr Phe Phe Arg Leu 1305 1310 1315 1320	4470
TTC CGA GTG ATG CGA TTG GTG AAG CTT CTC AGC AGG GGG GAA GGC ATC Phe Arg Val Met Arg Leu Val Lys Leu Leu Ser Arg Gly Glu Gly Ile 1325 1330 1335	4518
CGG ACA TTG CTG TGG ACT TTT ATT AAG TTC TTT CAG GCG CTC CCG TAT Arg Thr Leu Leu Trp Thr Phe Ile Lys Phe Phe Gln Ala Leu Pro Tyr 1340 1345 1350	4566
GTG GCC CTC CTC ATA GCC ATG CTG TTC TTC ATC TAT GCG GTC ATT GGC Val Ala Leu Leu Ile Ala Met Leu Phe Phe Ile Tyr Ala Val Ile Gly 1355 1360 1365	4614
ATG CAG ATG TTT GGG AAA GTT GCC ATG AGA GAT AAC AAC CAG ATC AAT Met Gln Met Phe Gly Lys Val Ala Met Arg Asp Asn Asn Gln Ile Asn 1370 1375 1380	4662
AGG AAC AAT AAC TTC CAG ACG TTT CCC CAG GCG GTG CTG CTG CTC TTC Arg Asn Asn Asn Phe Gln Thr Phe Pro Gln Ala Val Leu Leu Phe 1385 1390 1395 1400	4710
AGG TGT GCA ACA GGT GAG GCC TGG CAG GAG ATC ATG CTG GCC TGT CTC Arg Cys Ala Thr Gly Glu Ala Trp Gln Glu Ile Met Leu Ala Cys Leu 1405 1410 1415	4758
CCA GGG AAG CTC TGT GAC CCT GAG TCA GAT TAC AAC CCC GGG GAG GAG Pro Gly Lys Leu Cys Asp Pro Glu Ser Asp Tyr Asn Pro Gly Glu Glu 1420 1425 1430	4806
CAT ACA TGT GGG AGC AAC TTT GCC ATT GTC TAT TTC ATC AGT TTT TAC His Thr Cys Gly Ser Asn Phe Ala Ile Val Tyr Phe Ile Ser Phe Tyr 1435 1440 1445	4854

-99-

ATG CTC TGT GCA TTT CTG ATC ATC AAT CTG TTT GTG GCT GTC ATC ATG Met Leu Cys Ala Phe Leu Ile Ile Asn Leu Phe Val Ala Val Ile Met 1450 1455 1460	4902
GAT AAT TTC GAC TAT CTG ACC CGG GAC TGG TCT ATT TTG GGG CCT CAC Asp Asn Phe Asp Tyr Leu Thr Arg Asp Trp Ser Ile Leu Gly Pro His 1465 1470 1475 1480	4950
CAT TTA GAT GAA TTC AAA AGA ATA TGG TCA GAA TAT GAC CCT GAG GCA His Leu Asp Glu Phe Lys Arg Ile Trp Ser Glu Tyr Asp Pro Glu Ala 1485 1490 1495	4998
AAG GGA AGG ATA AAA CAC CTT GAT GTG GTC ACT CTG CTT CGA CGC ATC Lys Gly Arg Ile Lys His Leu Asp Val Val Thr Leu Leu Arg Arg Ile 1500 1505 1510	5046
CAG CCT CCC CTG GGG TTT GGG AAG TTA TGT CCA CAC AGG GTA GCG TGC Gln Pro Pro Leu Gly Phe Gly Lys Leu Cys Pro His Arg Val Ala Cys 1515 1520 1525	5094
AAG AGA TTA GTT GCC ATG AAC ATG CCT CTC AAC AGT GAC GGG ACA GTC Lys Arg Leu Val Ala Met Asn Met Pro Leu Asn Ser Asp Gly Thr Val 1530 1535 1540	5142
ATG TTT AAT GCA ACC CTG TTT GCT TTG GTT CGA ACG GCT CTT AAG ATC Met Phe Asn Ala Thr Leu Phe Ala Leu Val Arg Thr Ala Leu Lys Ile 1545 1550 1555 1560	5190
AAG ACC GAA GGG AAC CTG GAG CAA GCT AAT GAA GAA CTT CGG GCT GTG Lys Thr Glu Gly Asn Leu Glu Gln Ala Asn Glu Glu Leu Arg Ala Val 1565 1570 1575	5238
ATA AAG AAA ATT TGG AAG AAA ACC AGC ATG AAA TTA CTT GAC CAA GTT Ile Lys Ile Trp Lys Thr Ser Met Lys Leu Leu Asp Gln Val 1580 1585 1590	5286
GTC CCT CCA GCT GGT GAT GAT GAG GTA ACC GTG GGG AAG TTC TAT GCC Val Pro Pro Ala Gly Asp Asp Glu Val Thr Val Gly Lys Phe Tyr Ala 1595 1600 1605	5334
ACT TTC CTG ATA CAG GAC TAC TTT AGG AAA TTC AAG AAA CGG AAA GAA Thr Phe Leu Ile Gln Asp Tyr Phe Arg Lys Phe Lys Lys Arg Lys Glu 1610 1615 1620	5382
CAA GGA CTG GTG GGA AAG TAC CCT GCG AAG AAC ACC ACA ATT GCC CTA Gln Gly Leu Val Gly Lys Tyr Pro Ala Lys Asn Thr Thr Ile Ala Leu 1625 1630 1635 1640	5430
CAG GCG GGA TTA AGG ACA CTG CAT GAC ATT GGG CCA GAA ATC CGG CGT Gln Ala Gly Leu Arg Thr Leu His Asp Ile Gly Pro Glu Ile Arg Arg 1645 1650 1655	5478
GCT ATA TCG TGT GAT TTG CAA GAT GAC GAG CCT GAG GAA ACA AAA CGA Ala Ile Ser Cys Asp Leu Gln Asp Asp Glu Pro Glu Glu Thr Lys Arg 1660 1665 1670	5526
GAA GAA GAA GAT GAT GTG TTC AAA AGA AAT GGT GCC CTG CTT GGA AAC Glu Glu Glu Asp Asp Val Phe Lys Arg Asn Gly Ala Leu Leu Gly Asn 1675 1680 1685	5574
CAT GTC AAT CAT GTT AAT AGT GAT AGG AGA GAT TCC CTT CAG CGG ACC His Val Asn His Val Asn Ser Asp Arg Arg Asp Ser Leu Gln Gln Thr 1690 1695 1700	5622
AAT ACC ACC CAC CGT CCC CTG CAT GTC CAA AGG CCT TCA ATT CCA CCT Asn Thr Thr His Arg Pro Leu His Val Gln Arg Pro Ser Ile Pro Pro 1705 1710 1715 1720	5670

-100-

GCA AGT GAT ACT GAG AAA CCG CTG TTT CCT CCA GCA GGA AAT TCG GTG Ala Ser Asp Thr Glu Lys Pro Leu Phe Pro Pro Ala Gly Asn Ser Val 1725 1730 1735	5718
TGT CAT AAC CAT CAT AAC CAT AAT TCC ATA GGA AAG CAA GTT CCC ACC Cys His Asn His His Asn Ser Ile Gly Lys Gln Val Pro Thr 1740 1745 1750	5766
TCA ACA AAT GCC AAT CTC AAT AAT GCC AAT ATG TCC AAA GCT GCC CAT Ser Thr Asn Ala Asn Leu Asn Ala Asn Met Ser Lys Ala Ala His 1755 1760 1765	5814
GGA AAG CGG CCC AGC ATT GGG AAC CTT GAG CAT GTG TCT GAA AAT GGG Gly Lys Arg Pro Ser Ile Gly Asn Leu Glu His Val Ser Glu Asn Gly 1770 1775 1780	5862
CAT CAT TCT TCC CAC AAG CAT GAC CGG GAG CCT CAG AGA AGG TCC AGT His His Ser Ser His Lys His Asp Arg Glu Pro Gln Arg Arg Ser Ser 1785 1790 1795 1800	5910
GTG AAA AGA ACC CGC TAT TAT GAA ACT TAC ATT AGG TCC GAC TCA GGA Val Lys Arg Thr Arg Tyr Tyr Glu Thr Tyr Ile Arg Ser Asp Ser Gly 1805 1810 1815	5958
GAT GAA CAG CTC CCA ACT ATT TGC CGG GAA GAC CCA GAG ATA CAT GGC Asp Glu Gln Leu Pro Thr Ile Cys Arg Glu Asp Pro Glu Ile His Gly 1820 1825 1830	6006
TAT TTC AGG GAC CCC CAC TGC TTG GGG GAG CAG GAG TAT TTC AGT AGT Tyr Phe Arg Asp Pro His Cys Leu Gly Glu Gln Glu Tyr Phe Ser Ser 1835 1840 1845	6054
GAG GAA TGC TAC GAG GAT GAC AGC TCG CCC ACC TGG AGC AGG CAA AAC Glu Glu Cys Tyr Glu Asp Asp Ser Ser Pro Thr Trp Ser Arg Gln Asn 1850 1855 1860	6102
TAT GGC TAC TAC AGC AGA TAC CCA GGC AGA AAC ATC GAC TCT GAG AGG Tyr Gly Tyr Tyr Ser Arg Tyr Pro Gly Arg Asn Ile Asp Ser Glu Arg 1865 1870 1875 1880	6150
CCC CGA GGC TAC CAT CAT CCC CAA GGA TTC TTG GAG GAC GAT GAC TCG Pro Arg Gly Tyr His His Pro Gln Gly Phe Leu Glu Asp Asp Ser 1885 1890 1895	6198
CCC GTT TGC TAT GAT TCA CGG AGA TCT CCA AGG AGA CGC CTA CTA CCT Pro Val Cys Tyr Asp Ser Arg Arg Ser Pro Arg Arg Arg Leu Leu Pro 1900 1905 1910	6246
CCC ACC CCA GCA TCC CAC CGG AGA TCC TCC TTC AAC TTT GAG TGC CTG Pro Thr Pro Ala Ser His Arg Arg Ser Ser Phe Asn Phe Glu Cys Leu 1915 1920 1925	6294
CGC CGG CAG AGC AGC CAG GAA GAG GTC CCG TCG TCT CCC ATC TTC CCC Arg Arg Gln Ser Ser Gln Glu Glu Val Pro Ser Ser Pro Ile Phe Pro 1930 1935 1940	6342
CAT CGC ACG GCC CTG CCT CTG CAT CTA ATG CAG CAA CAG ATC ATG GCA His Arg Thr Ala Leu Pro Leu His Leu Met Gln Gln Gln Ile Met Ala 1945 1950 1955 1960	6390
GTT GCC GGC CTA GAT TCA AGT AAA GCC CAG AAG TAC TCA CCG AGT CAC Val Ala Gly Leu Asp Ser Ser Lys Ala Gln Lys Tyr Ser Pro Ser His 1965 1970 1975	6438
TCG ACC CGG TCG TGG GCC ACC CCT CCA GCA ACC CCT CCC TAC CGG GAC Ser Thr Arg Ser Trp Ala Thr Pro Pro Ala Thr Pro Pro Tyr Arg Asp 1980 1985 1990	6486

-101-

TGG ACA CCG TGC TAC ACC CCC CTG ATC CAA GTG GAG CAG TCA GAG GCC Trp Thr Pro Cys Tyr Thr Pro Leu Ile Gln Val Glu Gln Ser Glu Ala 1995 2000 2005	6534
CTG GAC CAG GTG AAC GGC AGC CTG CCG TCC CTG CAC CGC AGC TCC TGG Leu Asp Gln Val Asn Gly Ser Leu Pro Ser Leu His Arg Ser Ser Trp 2010 2015 2020	6582
TAC ACA GAC GAG CCC GAC ATC TCC TAC CCG ACT TTC ACA CCA GCC AGC Tyr Thr Asp Glu Pro Asp Ile Ser Tyr Arg Thr Phe Thr Pro Ala Ser 2025 2030 2035 2040	6630
CTG ACT GTC CCC AGC AGC TTC CCG AAC AAA AAC AGC GAC AAG CAG AGG Leu Thr Val Pro Ser Ser Phe Arg Asn Lys Asn Ser Asp Lys Gln Arg 2045 2050 2055	6678
AGT GCG GAC AGC TTG GTG GAG GCA GTC CTG ATA TCC GAA GGC TTG GGA Ser Ala Asp Ser Leu Val Glu Ala Val Leu Ile Ser Glu Gly Leu Gly 2060 2065 2070	6726
CGC TAT GCA AGG GAC CCA AAA TTT GTG TCA GCA ACA AAA CAC GAA ATC Arg Tyr Ala Arg Asp Pro Lys Phe Val Ser Ala Thr Lys His Glu Ile 2075 2080 2085	6774
GCT GAT GCC TGT GAC CTC ACC ATC GAC GAG ATG GAG AGT GCA GCC AGC Ala Asp Ala Cys Asp Leu Thr Ile Asp Glu Met Glu Ser Ala Ala Ser 2090 2095 2100	6822
ACC CTG CTT AAT GGG AAC GTG CGT CCC CGA GCC AAC GGG GAT GTG GGC Thr Leu Leu Asn Gly Asn Val Arg Pro Arg Ala Asn Gly Asp Val Gly 2105 2110 2115 2120	6870
CCC CTC TCA CAC CGG CAG GAC TAT GAG CTA CAG GAC TTT GGT CCT GGC Pro Leu Ser His Arg Gln Asp Tyr Glu Leu Gln Asp Phe Gly Pro Gly 2125 2130 2135	6918
TAC AGC GAC GAA GAG CCA GAC CCT GGG AGG GAT GAG GAG GAC CTG GCG Tyr Ser Asp Glu Glu Pro Asp Pro Gly Arg Asp Glu Glu Asp Leu Ala 2140 2145 2150	6966
GAT GAA ATG ATA TGC ATC ACC ACC TTG TAGCCCCAG CGAGGGCAG Asp Glu Met Ile Cys Ile Thr Thr Leu 2155 2160	7013
ACTGGCTCTG GCCTCAGGTG GGGCGCAGGA GAGCCAGGG AAAAGTGCCT CATAGTTAGG AAAGTTAGG CACTAGTTGG GAGTAATATT CAATTAATTA GACTTTGTGTA TAAGAGATGT	7073
CATGCCTCAA GAAAGCCATA AACCTGGTAG GAACAGGTCC CAAGCGGTTG AGCCTGGCAG	7133
AGTACCATGC GCTCGGCCCG AGCTGCAGGA AACAGCAGGC CCCGCCCTCT CACAGAGGAT	7193
GGGTGAGGAG GCCAGACCTG CCCTGCCCA TTGTCCAGAT GGGCACTGCT GTGGAGTCTG	7253
CTTCTCCCCT GTACCAGGGC ACCAGGCCCA CCCAACTGAA GGCATGGCGG CGGGGTGCAG	7313
GGGAAAGTTA AAGGTGATGA CGATCATCAC ACCTGTGTCG TTACCTCAGC CATCGGTCTA	7373
GCATATCAGT CACTGGGCC AACATATCCA TTTTAAACC CTTTCCCCCA AATACACTGC	7433
GTCCTGGTTC CTGTTAGCT GTTCTGAAAT ACGGTGTGTA AGTAAGTCAG AACCCAGCTA	7493
CCAGTGATTA TTGCGAGGGC AATGGGACCT CATAAATAAG GTTTCTGTG ATGTGACGCC	7553
AGTTTACATA AGAGAATATC AC	7613
	7635

-102-

## (2) INFORMATION FOR SEQ ID NO:2:

## (i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 104 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: double
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (genomic)

## (ix) FEATURE:

- (A) NAME/KEY: CDS
- (B) LOCATION: 1..102

## (ix) FEATURE:

- (A) NAME/KEY: misc\_feature
- (B) LOCATION: 1..104
- (D) OTHER INFORMATION: /note= "A 104-nucleotide alternative exon of alpha-1D."

## (xi) SEQUENCE DESCRIPTION: SEQ ID NO:2:

GTA AAT GAT GCG ATA GGA TGG GAA TGG CCA TGG GTG TAT TTT GTT AGT	48
Val Asn Asp Ala Ile Gly Trp Glu Trp Pro Trp Val Tyr Phe Val Ser	
1                   5                   10                   15	

CTG ATC ATC CTT GGC TCA TTT TTC GTC CTT AAC CTG GTT CTT GGT GTC	96
Leu Ile Ile Leu Gly Ser Phe Phe Val Leu Asn Leu Val Leu Gly Val	
20                 25                 30	

CTT AGT GG Leu Ser	104
-----------------------	-----

## (2) INFORMATION FOR SEQ ID NO:3:

## (i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 5904 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (genomic)

## (ix) FEATURE:

- (A) NAME/KEY: CDS
- (B) LOCATION: 1..5904

## (xi) SEQUENCE DESCRIPTION: SEQ ID NO:3:

ATG GTC AAT GAG AAT ACG AGG ATG TAC ATT CCA GAG GAA AAC CAC CAA	48
Met Val Asn Glu Asn Thr Arg Met Tyr Ile Pro Glu Glu Asn His Gln	
1                 5                 10                 15	

GGT TCC AAC TAT GGG AGC CCA CGC CCC GCC CAT GCC AAC ATG AAT GCC	96
Gly Ser Asn Tyr Gly Ser Pro Arg Pro Ala His Ala Asn Met Asn Ala	
20                 25                 30	

AAT GCG GCA GCG GGG CTG GCC CCT GAG CAC ATC CCC ACC CCG GGG GCT	144
Asn Ala Ala Ala Gly Leu Ala Pro Glu His Ile Pro Thr Pro Gly Ala	
35                 40                 45	

GCC CTG TCG TGG CAG GCG GCC ATC GAC GCA GCC CGG CAG GCT AAG CTG	192
Ala Leu Ser Trp Gln Ala Ala Ile Asp Ala Ala Arg Gln Ala Lys Leu	
50                 55                 60	

-103-

ATG GGC AGC GCT GGC AAT GCG ACC ATC TCC ACA GTC AGC TCC ACG CAG Met Gly Ser Ala Ile Asn Ala Thr Ile Ser Thr Val Ser Ser Thr Gln 65 70 75 80	240
CGG AAG CGC CAG CAA TAT GGG AAA CCC AAG AAG CAG GGC AGC ACC ACG Arg Lys Arg Gln Gln Tyr Gly Lys Pro Lys Lys Gln Gly Ser Thr Thr 85 90 95	288
GCC ACA CGC CCG CCC CGA GCC CTG CTC TGC CTG ACC CTG AAG AAC CCC Ala Thr Arg Pro Pro Arg Ala Leu Leu Cys Leu Thr Leu Lys Asn Pro 100 105 110	336
ATC CGG AGG GCC TGC ATC AGC ATT GTC GAA TGG AAA CCA TTT GAA ATA Ile Arg Arg Ala Cys Ile Ser Ile Val Glu Trp Lys Pro Phe Glu Ile 115 120 125	384
ATT ATT TTA CTG ACT ATT TTT GCC AAT TGT GTG GCC TTA GCG ATC TAT Ile Ile Leu Leu Thr Ile Phe Ala Asn Cys Val Ala Leu Ala Ile Tyr 130 135 140	432
ATT CCC TTT CCA GAA GAT GAT TCC AAC GCC ACC AAT TCC AAC CTG GAA Ile Pro Phe Pro Glu Asp Asp Ser Asn Ala Thr Asn Ser Asn Leu Glu 145 150 155 160	480
CGA GTG GAA TAT CTC TTT CTC ATA ATT TTT ACG GTG GAA GCG TTT TTA Arg Val Glu Tyr Leu Phe Leu Ile Ile Phe Thr Val Glu Ala Phe Leu 165 170 175	528
AAA GTA ATC GCC TAT GGA CTC CTC TTT CAC CCC AAT GCC TAC CTC CGC Lys Val Ile Ala Tyr Gly Leu Leu Phe His Pro Asn Ala Tyr Leu Arg 180 185 190	576
AAC GGC TGG AAC CTA CTA GAT TTT ATA ATT GTG GTT GTG GGG CTT TTT Asn Gly Trp Asn Leu Leu Asp Phe Ile Ile Val Val Val Gly Leu Phe 195 200 205	624
AGT GCA ATT TTA GAA CAA GCA ACC AAA GCA GAT GGG GCA AAC GCT CTC Ser Ala Ile Leu Glu Gln Ala Thr Lys Ala Asp Gly Ala Asn Ala Leu 210 215 220	672
GGA GGG AAA GGG GCC GGA TTT GAT GTG AAG GCG CTG AGG GCC TTC CGC Gly Gly Lys Gly Ala Gly Phe Asp Val Lys Ala Leu Arg Ala Phe Arg 225 230 235 240	720
GTG CTG CGC CCC CTG CGG CTG GTG TCC GGA GTC CCA AGT CTC CAG GTG Val Leu Arg Pro Leu Arg Leu Val Ser Gly Val Pro Ser Leu Gln Val 245 250 255	768
GTC CTG AAT TCC ATC ATC AAG GCC ATG GTC CCC CTG CTG CAC ATC GCC Val Leu Asn Ser Ile Ile Lys Ala Met Val Pro Leu Leu His Ile Ala 260 265 270	816
CTG CTT GTG CTG TTT GTC ATC ATC ATC TAC GCC ATC ATC GGC TTG GAG Leu Leu Val Leu Phe Val Ile Ile Ile Tyr Ala Ile Ile Gly Leu Glu 275 280 285	864
CTC TTC ATG GGG AAG ATG CAC AAG ACC TGC TAC AAC CAG GAG GGC ATA Leu Phe Met Gly Lys Met His Lys Thr Cys Tyr Asn Gln Glu Gly Ile 290 295 300	912
GCA GAT GTT CCA GCA GAA GAT GAC CCT TCC CCT TGT GCG CTG GAA ACG Ala Asp Val Pro Ala Glu Asp Asp Pro Ser Pro Cys Ala Leu Glu Thr 305 310 315 320	960
GGC CAC GGG CGG CAG TGC CAG AAC GGC ACG GTG TGC AAG CCC GGC TGG Gly His Gly Arg Gln Cys Gln Asn Gly Thr Val Cys Lys Pro Gly Trp 325 330 335	1008

-104-

GAT GGT CCC AAG CAC GGC ATC ACC AAC TTT GAC AAC TTT GCC TTC GCC Asp Gly Pro Lys His Gly Ile Thr Asn Phe Asp Asn Phe Ala Phe Ala 340 345 350	1056
ATG CTC ACG GTG TTC CAG TGC ATC ACC ATG GAG GGC TGG ACG GAC GTG Met Leu Thr Val Phe Gln Cys Ile Thr Met Glu Gly Trp Thr Asp Val 355 360 365	1104
CTG TAC TGG GTC AAT GAT GCC GTA GGA AGG GAC TGG CCC TGG ATC TAT Leu Tyr Trp Val Asn Asp Ala Val Gly Arg Asp Trp Pro Trp Ile Tyr 370 375 380	1152
TTT GTT ACA CTA ATC ATC ATA GGG TCA TTT TTT GTA CTT AAC TTG GTT Phe Val Thr Leu Ile Ile Gly Ser Phe Val Leu Asn Leu Val 385 390 395 400	1200
CTC GGT GTG CTT AGC GGA GAG TTT TCC AAA GAG AGG GAG AAG GCC AAG Leu Gly Val Leu Ser Gly Glu Phe Ser Lys Glu Arg Glu Lys Ala Lys 405 410 415	1248
GCC CGG GGA GAT TTC CAG AAG CTG CGG GAG AAG CAG CAG CTA GAA GAG Ala Arg Gly Asp Phe Gln Lys Leu Arg Glu Lys Gln Gln Leu Glu Glu 420 425 430	1296
GAT CTC AAA GGC TAC CTG GAT TGG ATC ACT CAG GCC GAA GAC ATC GNT Asp Leu Lys Gly Tyr Leu Asp Trp Ile Thr Gln Ala Glu Asp Ile Xaa 435 440 445	1344
CCT GAG AAT GAG GAC GAA GGC ATG GAT GAG GAG AAG CCC CGA AAC AGA Pro Glu Asn Glu Asp Glu Gly Met Asp Glu Glu Lys Pro Arg Asn Arg 450 455 460	1392
GGC ACT CCG GCG GGC ATG CTT GAT CAG AAG AAA GGG AAG TTT GCT TGG Gly Thr Pro Ala Gly Met Leu Asp Gln Lys Lys Gly Lys Phe Ala Trp 465 470 475 480	1440
TTT AGT CAC TCC ACA GAA ACC CAT GTG AGC ATG CCC ACC AGT GAG ACC Phe Ser His Ser Thr Glu Thr His Val Ser Met Pro Thr Ser Glu Thr 485 490 495	1488
GAG TCC GTC AAC ACC GAA AAC GTG GCT GGA GGT GAC ATC GAG GGA GAA Glu Ser Val Asn Thr Glu Asn Val Ala Gly Gly Asp Ile Glu Gly Glu 500 505 510	1536
AAC TGC GGG GCC AGG CTG GCC CAC CGG ATC TCC AAG TCA AAG TTC AGC Asn Cys Gly Ala Arg Leu Ala His Arg Ile Ser Lys Ser Lys Phe Ser 515 520 525	1584
CGC TAC TGG CGC CGG TGG AAT CGG TTC TGC AGA AGG AAG TGC CGC GCC Arg Tyr Trp Arg Arg Trp Asn Arg Phe Cys Arg Arg Lys Cys Arg Ala 530 535 540	1632
GCA GTC AAG TCT AAT GTC TTC TAC TGG CTG GTG ATT TTC CTG GTG TTC Ala Val Lys Ser Asn Val Phe Tyr Trp Leu Val Ile Phe Leu Val Phe 545 550 555 560	1680
CTC AAC ACG CTC ACC ATT GCC TCT GAG CAC TAC AAC CAG CCC AAC TGG Leu Asn Thr Leu Thr Ile Ala Ser Glu His Tyr Asn Gln Pro Asn Trp 565 570 575	1728
CTC ACA GAA GTC CAA GAC ACG GCA AAC AAG GCC CTG CTG GCC CTG TTC Leu Thr Glu Val Gln Asp Thr Ala Asn Lys Ala Leu Leu Ala Leu Phe 580 585 590	1776
ACG GCA GAG ATG CTC CTG AAG ATG TAC AGC CTG GGC CTG CAG GCC TAC Thr Ala Glu Met Leu Leu Lys Met Tyr Ser Leu Gly Leu Gln Ala Tyr 595 600 605	1824

-105-

TTC GTG TCC CTC TTC AAC CGC TTT GAC TGC TTC GTC GTG TGT GGC GGC Phe Val Ser Leu Phe Asn Arg Phe Asp Cys Phe Val Val Cys Gly Gly 610 615 620	1872
ATC CTG GAG ACC ATC CTG GTG GAG ACC AAG ATC ATG TCC CCA CTG GGC Ile Leu Glu Thr Ile Leu Val Glu Thr Lys Ile Met Ser Pro Leu Gly 625 630 635 640	1920
ATC TCC GTG CTC AGA TGC GTC CGG CTG CTG AGG ATT TTC AAG ATC ACG Ile Ser Val Leu Arg Cys Val Arg Leu Leu Arg Ile Phe Lys Ile Thr 645 650 655	1968
AGG TAC TGG AAC TCC TTG AGC AAC CTG GTG GCA TCC TTG CTG AAC TCT Arg Tyr Trp Asn Ser Leu Ser Asn Leu Val Ala Ser Leu Leu Asn Ser 660 665 670	2016
GTC CGC TCC ATC GCC TCC CTG CTC CTT CTC CTC TTC CTC TTC ATC ATC Val Arg Ser Ile Ala Ser Leu Leu Leu Leu Phe Leu Phe Ile Ile 675 680 685	2064
ATC TTC TCC CTC CTG GGG ATG CAG CTC TTT GGA GGA AAG TTC AAC TTT Ile Phe Ser Leu Leu Gly Met Gln Leu Phe Gly Gly Lys Phe Asn Phe 690 695 700	2112
GAT GAG ATG CAG ACC CGG AGG AGC ACA TTC GAT AAC TTC CCC CAG TCC Asp Glu Met Gln Thr Arg Arg Ser Thr Phe Asp Asn Phe Pro Gln Ser 705 710 715 720	2160
CTC CTC ACT GTG TTT CAG ATC CTG ACC GGG GAG GAC TGG AAT TCG GTG Leu Leu Thr Val Phe Gln Ile Leu Thr Gly Glu Asp Trp Asn Ser Val 725 730 735	2208
ATG TAT GAT GGG ATC ATG GCT TAT GGG GGC CCC TCT TTT CCA GGG ATG Met Tyr Asp Gly Ile Met Ala Tyr Gly Gly Pro Ser Phe Pro Gly Met 740 745 750	2256
TTA GTC TGT ATT TAC TTC ATC ATC CTC TTC ATC TCT GGA AAC TAT ATC Leu Val Cys Ile Tyr Phe Ile Ile Leu Phe Ile Ser Gly Asn Tyr Ile 755 760 765	2304
CTA CTG AAT GTG TTC TTG GCC ATT GCT GTG GAC AAC CTG GCT GAT GCT Leu Leu Asn Val Phe Leu Ala Ile Ala Val Asp Asn Leu Ala Asp Ala 770 775 780	2352
GAG AGC CTC ACA TCT GCC CTA AAG GAG GAG GAA GAG GAG AAG GAG AGA Glu Ser Leu Thr Ser Ala Leu Lys Glu Glu Glu Glu Lys Glu Arg 785 790 795 800	2400
AAG AAG CTG GCC AGG ACT GCC AGC CCA GAG AAG AAA CAA GAG TTG GTG Lys Lys Leu Ala Arg Thr Ala Ser Pro Glu Lys Lys Gln Glu Leu Val 805 810 815	2448
GAG AAG CCG GCA GTG GGG GAA TCC AAG GAG GAG AAG ATT GAG CTG AAA Glu Lys Pro Ala Val Gly Glu Ser Lys Glu Glu Lys Ile Glu Leu Lys 820 825 830	2496
TCC ATC ACG GCT GAC GGA GAG TCT CCA CCC GCC ACC AAG ATC AAC ATG Ser Ile Thr Ala Asp Gly Glu Ser Pro Pro Ala Thr Lys Ile Asn Met 835 840 845	2544
GAT GAC CTC CAG CCC AAT GAA AAT GAG GAT ARG AGC CCC TAC CCC AAC Asp Asp Leu Gln Pro Asn Glu Asn Ile Asp Lys Ser Pro Tyr Pro Asn 850 855 860	2592
CCA GAA ACT ACA GGA GAA GAG GAT GAG GAG GAG CCA GAG ATG CCT GTC Pro Glu Thr Thr Gly Glu Glu Asp Glu Glu Glu Pro Glu Met Pro Val 865 870 875 880	2640

-106-

GGC CCT CGC CCA CGA CCA CTC TCT GAG CTT CAC CTT AAG GAA AAG GCA Gly Pro Arg Pro Arg Pro Leu Ser Glu Leu His Leu Lys Glu Lys Ala 885 890 895	2688
GTG CCC ATG CCA GAA GCC AGC GCG TTT TTC ATC TTC AGC TCT AAC AAC Val Pro Met Pro Glu Ala Ser Ala Phe Phe Ile Phe Ser Ser Asn Asn 900 905 910	2736
AGG TTT CGC CTC CAG TGC CAC CGC ATT GTC AAT GAC ACG ATC TTC ACC Arg Phe Arg Leu Gln Cys His Arg Ile Val Asn Asp Thr Ile Phe Thr 915 920 925	2784
AAC CTG ATC CTC TTC TTC ATT CTG CTC AGC AGC ATT TCC CTG GCT GCT Asn Leu Ile Leu Phe Phe Ile Leu Leu Ser Ser Ile Ser Leu Ala Ala 930 935 940	2832
GAG GAC CCG GTC CAG CAC ACC TCC TTC AGG AAC CAT ATT CTG TTT TAT Glu Asp Pro Val Gln His Thr Ser Phe Arg Asn His Ile Leu Phe Tyr 945 950 955 960	2880
TTT GAT ATT GTT TTT ACC ACC ATT TTC ACC ATT GAA ATT GCT CTG AAG Phe Asp Ile Val Phe Thr Thr Ile Phe Thr Ile Glu Ile Ala Leu Lys 965 970 975	2928
ATG ACT GCT TAT GGG GCT TTC TTG CAC AAG GGT TCT TTC TGC CGG AAC Met Thr Ala Tyr Gly Ala Phe Leu His Lys Gly Ser Phe Cys Arg Asn 980 985 990	2976
TAC TTC AAC ATC CTG GAC CTG CTG GTG GTC AGC GTG TCC CTC ATC TCC Tyr Phe Asn Ile Leu Asp Leu Leu Val Val Ser Val Ser Leu Ile Ser 995 1000 1005	3024
TTT GGC ATC CAG TCC AGT GCA ATC AAT GTC GTG AAG ATC TTG CGA GTC Phe Gly Ile Gln Ser Ser Ala Ile Asn Val Val Lys Ile Leu Arg Val 1010 1015 1020	3072
CTG CGA GTA CTC AGG CCC CTG AGG GCC ATC AAC AGG GCC AAG GGG CTA Leu Arg Val Leu Arg Pro Leu Arg Ala Ile Asn Arg Ala Lys Gly Leu 1025 1030 1035 1040	3120
AAG CAT GTG GTT CAG TGT GTG TTT GTC GCC ATC CGG ACC ATC GGG AAC Lys His Val Val Gln Cys Val Phe Val Ala Ile Arg Thr Ile Gly Asn 1045 1050 1055	3168
ATC GTG ATT GTC ACC ACC CTG CTG CAG TTC ATG TTT GCC TGC ATC GGG Ile Val Ile Val Thr Thr Leu Leu Gln Phe Met Phe Ala Cys Ile Gly 1060 1065 1070	3216
GTC CAG CTC TTC AAG GGA AAG CTG TAC ACC TGT TCA GAC AGT TCC AAG Val Gln Leu Phe Lys Gly Lys Leu Tyr Thr Cys Ser Asp Ser Ser Lys 1075 1080 1085	3264
CAG ACA GAG GCG GAA TGC AAG GGC AAC TAC ATC ACG TAC AAA GAC GGG Gln Thr Glu Ala Glu Cys Lys Gly Asn Tyr Ile Thr Tyr Lys Asp Gly 1090 1095 1100	3312
GAG GTT GAC CAC CCC ATC ATC CAA CCC CGC AGC TGG GAG AAC AGC AAG Glu Val Asp His Pro Ile Ile Gln Pro Arg Ser Trp Glu Asn Ser Lys 1105 1110 1115 1120	3360
TTT GAC TTT GAC AAT GTT CTG GCA GCC ATG ATG GCC CTC TTC ACC GTC Phe Asp Phe Asp Asn Val Leu Ala Ala Met Met Ala Leu Phe Thr Val 1125 1130 1135	3408
TCC ACC TTC GAA GGG TGG CCA GAG CTG CTG TAC CGC TCC ATC GAC TCC Ser Thr Phe Glu Gly Trp Pro Glu Leu Leu Tyr Arg Ser Ile Asp Ser 1140 1145 1150	3456

-107-

CAC ACG GAA GAC AAG GGC CCC ATC TAC AAC TAC CGT GTG GAG ATC TCC His Thr Glu Asp Lys Gly Pro Ile Tyr Asn Tyr Arg Val Glu Ile Ser 1155 1160 1165	3504
ATC TTC TTC ATC ATC TAC ATC ATC ATC ATC GCC TTC TTC ATG ATG AAC Ile Phe Phe Ile Ile Tyr Ile Ile Ile Ala Phe Phe Met Met Asn 1170 1175 1180	3552
ATC TTC GTG GGC TTC GTC ATC GTC ACC TTT CAG GAG CAG GGG GAG CAG Ile Phe Val Gly Phe Val Ile Val Thr Phe Gln Glu Gln Gly Glu Gln 1185 1190 1195 1200	3600
GAG TAC AAG AAC TGT GAG CTG GAC AAG AAC CAG CGA CAG TGC GTG GAA Glu Tyr Lys Asn Cys Glu Leu Asp Lys Asn Gln Arg Gln Cys Val Glu 1205 1210 1215	3648
TAC GCC CTC AAG GCC CGG CCC CTG CGG AGG TAC ATC CCC AAG AAC CAG Tyr Ala Leu Lys Ala Arg Pro Leu Arg Arg Tyr Ile Pro Lys Asn Gln 1220 1225 1230	3696
CAC CAG TAC AAA GTG TGG TAC GTG GTC AAC TCC ACC TAC TTC GAG TAC His Gln Tyr Lys Val Trp Tyr Val Val Asn Ser Thr Tyr Phe Glu Tyr 1235 1240 1245	3744
CTG ATG TTC GTC CTC ATC CTG CTC AAC ACC ATC TGC CTG GCC ATG CAG Leu Met Phe Val Leu Ile Leu Leu Asn Thr Ile Cys Leu Ala Met Gln 1250 1255 1260	3792
CAC TAC GGC CAG AGC TGC CTG TTC AAA ATC GCC ATG AAC ATC CTC AAC His Tyr Gly Gln Ser Cys Leu Phe Lys Ile Ala Met Asn Ile Leu Asn 1265 1270 1275 1280	3840
ATG CTC TTC ACT GGC CTC TTC ACC GTG GAG ATG ATC CTG AAG CTC ATT Met Leu Phe Thr Gly Leu Phe Thr Val Glu Met Ile Leu Lys Leu Ile 1285 1290 1295	3888
GCC TTC AAA CCC AAG GGT TAC TTT AGT GAT CCC TGG AAT GTT TTT GAC Ala Phe Lys Pro Lys Gly Tyr Phe Ser Asp Pro Trp Asn Val Phe Asp 1300 1305 1310	3936
TTC CTC ATC GTA ATT GGC AGC ATA ATT GAC GTC ATT CTC AGT GAG ACT Phe Leu Ile Val Ile Gly Ser Ile Ile Asp Val Ile Leu Ser Glu Thr 1315 1320 1325	3984
AAT CCA GCT GAA CAT ACC CAA TGC TCT CCC TCT ATG AAC GCA GAG GAA Asn Pro Ala Glu His Thr Gln Cys Ser Pro Ser Met Asn Ala Glu Glu 1330 1335 1340	4032
AAC TCC CGC ATC TCC ATC ACC TTC TTC CGC CTG TTC CGG GTC ATG CGT Asn Ser Arg Ile Ser Ile Thr Phe Phe Arg Leu Phe Arg Val Met Arg 1345 1350 1355 1360	4080
CTG GTG AAG CTG CTG AGC CGT GGG GAG GGC ATC CGG ACG CTG CTG TGG Leu Val Lys Leu Leu Ser Arg Gly Glu Gly Ile Arg Thr Leu Leu Trp 1365 1370 1375	4128
ACC TTC ATC AAG TCC TTC CAG GCC CTG CCC TAT GTG GCC CTC CTG ATC Thr Phe Ile Lys Ser Phe Gln Ala Leu Pro Tyr Val Ala Leu Leu Ile 1380 1385 1390	4176
GTG ATG CTG TTC ATC TAC GCG GTG ATC GGG ATG CAG GTG TTT GGG Val Met Leu Phe Phe Ile Tyr Ala Val Ile Gly Met Gln Val Phe Gly 1395 1400 1405	4224
AAA ATT GCC CTG AAT GAT ACC ACA GAG ATC AAC CGG AAC AAC AAC TTT Lys Ile Ala Leu Asn Asp Thr Thr Glu Ile Asn Arg Asn Asn Asn Phe 1410 1415 1420	4272

-108-

CAG ACC TTC CCC CAG GCC GTG CTG CTC CTC TTC AGG TGT GCC ACC GGG Gln Thr Phe Pro Gln Ala Val Leu Leu Phe Arg Cys Ala Thr Gly 1425 1430 1435 1440	4320
GAG GCC TGG CAG GAC ATC ATG CTG GCC TGC ATG CCA GGC AAG AAG TGT Glu Ala Trp Gln Asp Ile Met Leu Ala Cys Met Pro Gly Lys Lys Cys 1445 1450 1455	4368
GCC CCA GAG TCC GAG CCC AGC AAC AGC ACG GAG GGT GAA ACA CCC TGT Ala Pro Glu Ser Glu Pro Ser Asn Ser Thr Glu Gly Glu Thr Pro Cys 1460 1465 1470	4416
GGT AGC AGC TTT GCT GTC TTC TAC TTC ATC AGC TTC TAC ATG CGC TGT Gly Ser Ser Phe Ala Val Phe Tyr Phe Ile Ser Phe Tyr Met Arg Cys 1475 1480 1485	4464
GCC TTC CTG ATC ATC AAC CTC TTT GTA GCT GTC ATC ATG GAC AAC TTT Ala Phe Leu Ile Ile Asn Leu Phe Val Ala Val Ile Met Asp Asn Phe 1490 1495 1500	4512
GAC TAC CTG ACA AGG GAC TGG TCC ATC CTT GGT CCC CAC CAC CTG GAT Asp Tyr Leu Thr Arg Asp Trp Ser Ile Leu Gly Pro His His Leu Asp 1505 1510 1515 1520	4560
GAG TTT AAA AGA ATC TGG GCA GAG TAT GAC CCT GAA GCC AAG GGT CGT Glu Phe Lys Arg Ile Trp Ala Glu Tyr Asp Pro Glu Ala Lys Gly Arg 1525 1530 1535	4608
ATC AAA CAC CTG GAT GTG GTG ACC CTC CTC CGG CGG ATT CAG CCG CCA Ile Lys His Leu Asp Val Val Thr Leu Leu Arg Arg Ile Gln Pro Pro 1540 1545 1550	4656
CTA GGT TTT GGG AAG CTG TGC CCT CAC CGC GTG GCT TGC AAA CGC CTG Leu Gly Phe Gly Lys Leu Cys Pro His Arg Val Ala Cys Lys Arg Leu 1555 1560 1565	4704
GTC TCC ATG AAC ATG CCT CTG AAC AGC GAC GGG ACA GTC ATG TTC AAT Val Ser Met Asn Met Pro Leu Asn Ser Asp Gly Thr Val Met Phe Asn 1570 1575 1580	4752
GCC ACC CTG TTT GCC CTG GTC AGG ACG GCC CTG AGG ATC AAA ACA GAA Ala Thr Leu Phe Ala Leu Val Arg Thr Ala Leu Arg Ile Lys Thr Glu 1585 1590 1595 1600	4800
GGG AAC CTA GAA CAA GCC AAT GAG GAG CTG CGG GCG ATC ATC AAG AAG Gly Asn Leu Glu Gln Ala Asn Glu Glu Leu Arg Ala Ile Ile Lys Lys 1605 1610 1615	4848
ATC TGG AAG CGG ACC AGC ATG AAG CTG CTG GAC CAG GTG GTG CCC CCT Ile Trp Lys Arg Thr Ser Met Lys Leu Leu Asp Gln Val Val Pro Pro 1620 1625 1630	4896
GCA GGT GAT GAT GAG GTC ACC GTT GGC AAG TTC TAC GCC ACG TTC CTG Ala Gly Asp Asp Glu Val Thr Val Gly Lys Phe Tyr Ala Thr Phe Leu 1635 1640 1645	4944
ATC CAG GAG TAC TTC CGG AAG TTC AAG AAG CGC AAA GAG CAG GGC CTT Ile Gln Glu Tyr Phe Arg Lys Phe Lys Lys Arg Lys Glu Gln Gly Leu 1650 1655 1660	4992
GTG GGC AAG CCC TCC CAG AGG AAC GCG CTG TCT CTG CAG GCT GGC TTG Val Gly Lys Pro Ser Gln Arg Asn Ala Leu Ser Leu Gln Ala Gly Leu 1665 1670 1675 1680	5040
CGC ACA CTG CAT GAC ATC GGG CCT GAG ATC CGA CGG GCC ATC TCT GGA Arg Thr Leu His Asp Ile Gly Pro Glu Ile Arg Arg Ala Ile Ser Gly 1685 1690 1695	5088

-109-

GAT CTC ACC GCT GAG GAG GAG CTG GAC AAG GCC ATG AAG GAG GCT GTG Asp Leu Thr Ala Glu Glu Glu Leu Asp Lys Ala Met Lys Glu Ala Val 1700 1705 1710	5136
TCC GCT GCT TCT GAA GAT GAC ATC TTC AGG AGG GCC GGT GGC CTG TTC Ser Ala Ala Ser Glu Asp Asp Ile Phe Arg Arg Ala Gly Gly Leu Phe 1715 1720 1725	5184
GGC AAC CAC GTC AGC TAC TAC CAA AGC GAC GGC CGG AGC GCC TTC CCC Gly Asn His Val Ser Tyr Tyr Gln Ser Asp Gly Arg Ser Ala Phe Pro 1730 1735 1740	5232
CAG ACC TTC ACC ACT CAG CGC CCG CTG CAC ATC AAC AAG GCG GGC AGC Gln Thr Phe Thr Thr Gln Arg Pro Leu His Ile Asn Lys Ala Gly Ser 1745 1750 1755 1760	5280
AGC CAG GGC GAC ACT GAG TCG CCA TCC CAC GAG AAG CTG GTG GAC TCC Ser Gln Gly Asp Thr Glu Ser Pro Ser His Glu Lys Leu Val Asp Ser 1765 1770 1775	5328
ACC TTC ACC CCG AGC AGC TAC TCG TCC ACC GGC TCC AAC GCC AAC ATC Thr Phe Thr Pro Ser Ser Tyr Ser Ser Thr Gly Ser Asn Ala Asn Ile 1780 1785 1790	5376
AAC AAC GCC AAC AAC ACC GCC CTG GGT CGC CTC CCT CGC CCC GCC GGC Asn Asn Ala Asn Asn Thr Ala Leu Gly Arg Leu Pro Arg Pro Ala Gly 1795 1800 1805	5424
TAC CCC AGC ACA GTC AGC ACT GTG GAG GGC CAC GGG CCC CCC TTG TCC Tyr Pro Ser Thr Val Ser Thr Val Glu Gly His Gly Pro Pro Leu Ser 1810 1815 1820	5472
CCT GCC ATC CGG GTG CAG GAG GTG GCG TGG AAG CTC AGC TCC AAC AGG Pro Ala Ile Arg Val Gln Glu Val Ala Trp Lys Leu Ser Ser Asn Arg 1825 1830 1835 1840	5520
TGC CAC TCC CGG GAG AGC CAG GCA GCC ATG GCG CGT CAG GAG GAG ACG Cys His Ser Arg Glu Ser Gln Ala Ala Met Ala Arg Gln Glu Glu Thr 1845 1850 1855	5568
TCT CAG GAT GAG ACC TAT GAA GTG AAG ATG AAC CAT GAC ACG GAG GCC Ser Gln Asp Glu Thr Tyr Glu Val Lys Met Asn His Asp Thr Glu Ala 1860 1865 1870	5616
TGC AGT GAG CCC AGC CTG CTC TCC ACA GAG ATG CTC TCC TAC CAG GAT Cys Ser Glu Pro Ser Leu Leu Ser Thr Glu Met Leu Ser Tyr Gln Asp 1875 1880 1885	5664
GAC GAA AAT CGG CAA CTG ACG CTC CCA GAG GAG GAC AAG AGG GAC ATC Asp Glu Asn Arg Gln Leu Thr Leu Pro Glu Glu Asp Lys Arg Asp Ile 1890 1895 1900	5712
CGG CAA TCT CGG AAG AGG GGT TTC CTC CGC TCT TCC TCA CTA GGT CGA Arg Gln Ser Pro Lys Arg Gly Phe Leu Arg Ser Ser Ser Leu Gly Arg 1905 1910 1915 1920	5760
AGG GCC TCC TTC CAC CTG GAA TGT CTG AAG CGA CAG AAG GAC CGA GGG Arg Ala Ser Phe His Leu Glu Cys Leu Lys Arg Gln Lys Asp Arg Gly 1925 1930 1935	5808
GGA GAC ATC TCT CAG AAG ACA GTC CTG CCC TTG CAT CTG GTT CAT CAT Gly Asp Ile Ser Gln Lys Thr Val Leu Pro Leu His Leu Val His His 1940 1945 1950	5856
CAG GCA TTG GCA GTG GCA GGC CTG AGC CCC CTC CTC CAG AGA AGC CAT In Ala Leu Ala Val Ala Gly Leu Ser Pro Leu Leu Gln Arg Ser His 1955 1960 1965	5904

-110-

## (2) INFORMATION FOR SEQ ID NO:4:

## (i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 132 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: double
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (genomic)

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:4:

AGACCACGGC TTCCTCGAAT CTTGCGCGAA GCCGCCGGCC TCGGAGGAGG GATTAATCCA	60
GACCCGCCGG GGGGTGTTT CACATTCTT CCTCTTCGTG GCTGCTCCTC CTATTAAAAC	120
CATTTTG GT CC	132

## (2) INFORMATION FOR SEQ ID NO:5:

## (i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 89 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: double
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (genomic)

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:5:

CGCTGAGGGC CTTCCGCGTG CTGCGCCCCC TGCGGCTGGT GTCCGGAGTC CCAAGTCTCC	60
AGGTGGTCCT GAATTCCATC ATCAAGGCC	89

## (2) INFORMATION FOR SEQ ID NO:6:

## (i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 84 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: double
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (genomic)

(ix) FEATURE:

- (A) NAME/KEY: CDS
- (B) LOCATION: 1..84
- (D) OTHER INFORMATION: /note= "An alternative exon of alpha-1C."

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:6:

CAC TAT TTC TGT GAT GCA TGG AAT ACA TTT GAC GCC TTG ATT GTT GTG His Tyr Phe Cys Asp Ala Trp Asn Thr Phe Asp Ala Leu Ile Val Val	48
1 5 10 15	

GGT AGC ATT GTT GAT ATA GCA ATC ACC GAG GTA AAC Gly Ser Ile Val Asp Ile Ala Ile Thr Glu Val Asn	84
20 25	

## (2) INFORMATION FOR SEQ ID NO:7:

## (i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 7362 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: double
- (D) TOPOLOGY: linear

-111-

(ii) MOLECULE TYPE: DNA (genomic)

(ix) FEATURE:

- (A) NAME/KEY: CDS
- (B) LOCATION: 144..7163

(ix) FEATURE:

- (A) NAME/KEY: 5'UTR
- (B) LOCATION: 1..143

(ix) FEATURE:

- (A) NAME/KEY: 3'UTR
- (B) LOCATION: 7161..7362

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:7:

GC	GCGCCGGCGG	CTGCGGCCGGT	GGGGCCGGGC	GAGGTCCGTG	CGGTCCCCGGC	GGCTCCGTGG	60
CTG	CCTCCGCT	CTGAGCGCCT	GCGCGCCCCG	CGCCCTCCCT	GCCGGGGCCG	CTGGGGCCGGG	120
GAT	GCACGCG	GGGCCCGGGGA	GCC ATG GTC CGC TTC GGG GAC GAG CTG GGC	Met Val Arg Phe Gly Asp Glu Leu Gly			170
			1	5			
GGC	CGC TAT GGA GGC CCC GGC GGA GAG CGG GCC CGG GGC GGC GGG	Gly Arg Tyr Gly Pro Gly Gly Gly	Glu Arg Ala Arg Gly Gly Gly				218
10	15	20	25				
GCC	GGC GGG GCG GGG GGC CCG GGT CCC GGG GGG CTG CAG CCC GGC CAG	Ala Gly Ala Gly Pro Gly Pro Gly Gly	Lys Leu Gln Pro Gly Gln				266
30	35	40					
CGG	GTC CTC TAC AAG CAA TCG ATC GCG CAG CGC GCG CGG ACC ATG GCG	Arg Val Leu Tyr Lys Gln Ser Ile Ala Gln Arg Ala Arg Thr Met Ala					314
45	50	55					
CTG	TAC AAC CCC ATC CCG GTC AAG CAG AAC TGC TTC ACC GTC AAC CGC	Leu Tyr Asn Pro Ile Pro Val Lys Gln Asn Cys Phe Thr Val Asn Arg					362
60	65	70					
TCG	CTC TTC GTC TTC AGC GAG GAC AAC GTC GTC CGC AAA TAC GCG AAG	Ser Leu Phe Val Phe Ser Glu Asp Asn Val Val Arg Lys Tyr Ala Lys					410
75	80	85					
CGC	ATC ACC GAG TGG CCT CCA TTC GAG AAT ATG ATC CTG GCC ACC ATC	Arg Ile Thr Glu Trp Pro Pro Phe Glu Asn Met Ile Leu Ala Thr Ile					458
90	95	100	105				
ATC	GCC AAC TGC ATC GTG CTG GCC CTG GAG CAG CAC CTC CCT GAT GGG	Ile Ala Asn Cys Ile Val Leu Ala Leu Glu Gln His Leu Pro Asp Gly					506
110	115	120					
GAC	AAA ACG CCC ATG TCC GAG CGG CTG GAC GAC ACG GAG CCC TAT TTC	Asp Lys Thr Pro Met Ser Glu Arg Leu Asp Asp Thr Glu Pro Tyr Phe					554
125	130	135					
ATC	GGG ATC TTT TGC TTC GAG GCA GGG ATC AAA ATC ATC GCT CTG GGC	Ile Gly Ile Phe Cys Phe Glu Ala Gly Ile Lys Ile Ile Ala Leu Gly					602
140	145	150					
TTT	GTC TTC CAC AAG GGC TCT TAC CTG CGG AAC GGC TGG AAC GTC ATG	Phe Val Phe His Lys Gly Ser Tyr Leu Arg Asn Gly Trp Asn Val Met					650
155	160	165					
GAC	TTC GTG GTC CTC ACA GGG ATC CTT GCC ACG GCT GGA ACT GAC	Asp Phe Val Val Val Leu Thr Gly Ile Leu Ala Thr Ala Gly Thr Asp					698
170	175	180	185				

-112-

TTC GAC CTG CGA ACA CTG AGG GCT GTG CGT GTG CTG AGG CCC CTG AAG Phe Asp Leu Arg Thr Leu Arg Ala Val Arg Val Leu Arg Pro Leu Lys 190 195 200	746
CTG GTG TCT GGG ATT CCA AGT TTG CAG GTG GTG CTC AAG TCC ATC ATG Leu Val Ser Gly Ile Pro Ser Leu Gln Val Val Leu Lys Ser Ile Met 205 210 215	794
AAG GCC ATG GTT CCA CTC CTG CAG ATT GGG CTG CTT CTC TTC TTT GCC Lys Ala Met Val Pro Leu Leu Gln Ile Gly Leu Leu Leu Phe Phe Ala 220 225 230	842
ATC CTC ATG TTT GCC ATC ATT GGC CTG GAG TTC TAC ATG GGC AAG TTC Ile Leu Met Phe Ala Ile Ile Gly Leu Glu Phe Tyr Met Gly Lys Phe 235 240 245	890
CAC AAG GCC TGT TTC CCC AAC AGC ACA GAT GCG GAG CCC GTG GGT GAC His Lys Ala Cys Phe Pro Asn Ser Thr Asp Ala Glu Pro Val Gly Asp 250 255 260 265	938
TTC CCC TGT GGC AAG GAG GCC CCA GCC CGG CTG TGC GAG GGC GAC ACT Phe Pro Cys Gly Lys Glu Ala Pro Ala Arg Leu Cys Glu Gly Asp Thr 270 275 280	986
GAG TGC CGG GAG TAC TGG CCA GGA CCC AAC TTT GGC ATC ACC AAC TTT Glu Cys Arg Glu Tyr Trp Pro Gly Pro Asn Phe Gly Ile Thr Asn Phe 285 290 295	1034
GAC AAT ATC CTG TTT GCC ATC TTG ACG GTG TTC CAG TGC ATC ACC ACC ATG Asp Asn Ile Leu Phe Ala Ile Leu Thr Val Phe Gln Cys Ile Thr Met 300 305 310	1082
GAG GGC TGG ACT GAC ATC CTC TAT AAT ACA AAC GAT GCG GCC GGC AAC Glu Gly Trp Thr Asp Ile Leu Tyr Asn Thr Asn Asp Ala Ala Gly Asn 315 320 325	1130
ACC TGG AAC TGG CTC TAC TTC ATC CCT CTC ATC ATC ATC GGC TCC TTC Thr Trp Asn Trp Leu Tyr Phe Ile Pro Leu Ile Ile Ile Gly Ser Phe 330 335 340 345	1178
TTC ATG CTC AAC CTG GTG CTG GGC GTG CTC TCG GGG GAG TTT GCC AAG Phe Met Leu Asn Leu Val Leu Gly Val Leu Ser Gly Glu Phe Ala Lys 350 355 360	1226
GAG CGA GAG AGG GTG GAG AAC CGC CGC GCC TTC CTG AAG CTG CGC CGG Glu Arg Glu Arg Val Glu Asn Arg Arg Ala Phe Leu Lys Leu Arg Arg 365 370 375	1274
CAG CAG CAG ATC GAG CGA GAG CTC AAC GGG TAC CTG GAG TGG ATC TTC Gln Gln Gln Ile Glu Arg Glu Leu Asn Gly Tyr Leu Glu Trp Ile Phe 380 385 390	1322
AAG GCG GAG GAA GTC ATG CTG GCC GAG GAG GAC AGG AAT GCA GAG GAG Lys Ala Glu Glu Val Met Leu Ala Glu Glu Asp Arg Asn Ala Glu Glu 395 400 405	1370
AAG TCC CCT TTG GAC GTG CTG AAG AGA GCG GCC ACC AAG AAG AGC AGA Lys Ser Pro Leu Asp Val Leu Lys Arg Ala Ala Thr Lys Lys Ser Arg 410 415 420 425	1418
AAT GAC CTG ATC CAC GCA GAG GAG GGA GAG GAC CGG TTT GCA GAT CTC Asn Asp Leu Ile His Ala Glu Glu Gly Glu Asp Arg Phe Ala Asp Leu 430 435 440	1466
TGT GCT GTT GGA TCC CCC TTC GCC CGC GCC AGC CTC AAG AGC GGG AAG Cys Ala Val Gly Ser Pro Phe Ala Arg Ala Ser Leu Lys Ser Gly Lys 445 450 455	1514

-113-

ACA GAG AGC TCG TCA TAC TTC CCG AGG AAG GAG AAG ATG TTC CGG TTT Thr Glu Ser Ser Ser Tyr Phe Arg Arg Lys Glu Lys Met Phe Arg Phe 460 465 470	1562
TTT ATC CGG CGC ATG GTG AAG GCT CAG AGC TTC TAC TGG GTG GTG CTG Phe Ile Arg Arg Met Val Lys Ala Gln Ser Phe Tyr Trp Val Val Leu 475 480 485	1610
TGC GTG GTG GCC CTG AAC ACA CTG TGT GTG GCC ATG GTG CAT TAC AAC Cys Val Val Ala Leu Asn Thr Leu Cys Val Ala Met Val His Tyr Asn 490 495 500 505	1658
CAG CCG CGG CGG CTT ACC ACG ACC CTG TAT TTT GCA GAG TTT GTT TTC Gln Pro Arg Arg Leu Thr Thr Leu Tyr Phe Ala Glu Phe Val Phe 510 515 520	1706
CTG GGT CTC TTC CTC ACA GAG ATG TCC CTG AAG ATG TAT GGC CTG GGG Leu Gly Leu Leu Thr Glu Met Ser Leu Lys Met Tyr Gly Leu Gly 525 530 535	1754
CCC AGA AGC TAC TTC CGG TCC TCC TTC AAC TGC TTC GAC TTT GGG GTC Pro Arg Ser Tyr Phe Arg Ser Ser Phe Asn Cys Phe Asp Phe Gly Val 540 545 550	1802
ATC GTG GGG AGC GTC TTT GAA GTG GTC TGG GCG GCC ATC AAG CCG GGA Ile Val Gly Ser Val Phe Glu Val Val Trp Ala Ala Ile Lys Pro Gly 555 560 565	1850
AGC TCC TTT GGG ATC AGT GTG CTG CGG GCC CTC CGC CTG CTG AGG ATC Ser Ser Phe Gly Ile Ser Val Leu Arg Ala Leu Arg Leu Leu Arg Ile 570 575 580 585	1898
TTC AAA GTC ACG AAG TAC TGG AGC TCC CTG CGG AAC CTG GTG GTG TCC Phe Lys Val Thr Lys Tyr Trp Ser Ser Leu Arg Asn Leu Val Val Ser 590 595 600	1946
CTG CTG AAC TCC ATG AAG TCC ATC ATC AGC CTG CTC TTC TTG CTC TTC Leu Leu Asn Ser Met Lys Ser Ile Ile Ser Leu Leu Phe Leu Leu Phe 605 610 615	1994
CTG TTC ATT GTG GTC TTC GCC CTG CTG GGG ATG CAG CTG TTT GGG GGA Leu Phe Ile Val Val Phe Ala Leu Leu Gly Met Gln Leu Phe Gly Gly 620 625 630	2042
CAG TTC AAC TTC CAG GAT GAG ACT CCC ACA ACC AAC TTC GAC ACC TTC Gln Phe Asn Phe Gln Asp Glu Thr Pro Thr Thr Asn Phe Asp Thr Phe 635 640 645	2090
CCT GCC GCC ATC CTC ACT GTC TTC CAG ATC CTG ACN GGA GAG GAC TGG Pro Ala Ala Ile Leu Thr Val Phe Gln Ile Leu Thr Gly Glu Asp Trp 650 655 660 665	2138
AAT GCA GTG ATG TAT CAC GGG ATC GAA TCG CAA GGC GGC GTC AGC AAA Asn Ala Val Met Tyr His Gly Ile Glu Ser Gln Gly Gly Val Ser Lys 670 675 680	2186
GGC ATG TTC TCG TCC TTT TAC TTC ATT GTC CTG ACA CTG TTC GGA AAC Gly Met Phe Ser Ser Phe Tyr Phe Ile Val Leu Thr Leu Phe Gly Asn 685 690 695	2234
TAC ACT CTG CTG AAT GTC TTT CTG GCC ATC GCT GTG GAC AAC CTG GCC Tyr Thr Leu Leu Asn Val Phe Leu Ala Ile Ala Val Asp Asn Leu Ala 700 705 710	2282
AAC GCC CAA GAG CTG ACC AAG GAT GAA GAG GAG ATG GAA GAA GCA GCC Asn Ala Gln Glu Leu Thr Lys Asp Glu Glu Met Glu Glu Ala Ala 715 720 725	2330

-114-

AAT CAG AAG CTT GCT CTG CAA AAG GCC AAA GAA GTG GCT GAA GTC AGC Asn Gln Lys Leu Ala Leu Gln Lys Ala Lys Glu Val Ala Glu Val Ser 730 735 740 745	2378
CCC ATG TCT GCC GCG AAC ATC TCC ATC GCC GCC AGG CAG CAG AAC TCG Pro Met Ser Ala Ala Asn Ile Ser Ile Ala Ala Arg Gln Gln Asn Ser 750 755 760	2426
GCC AAG GCG CGC TCG GTG TGG GAG CAG CGG GCC AGC CAG CTA CGG CTG Ala Lys Ala Arg Ser Val Trp Glu Gln Arg Ala Ser Gln Leu Arg Leu 765 770 775	2474
CAG AAC CTG CGG GCC AGC TGC GAG GCG CTG TAC AGC GAG ATG GAC CCC Gln Asn Leu Arg Ala Ser Cys Glu Ala Leu Tyr Ser Glu Met Asp Pro 780 785 790	2522
GAG GAG CGG CTG CGC TTC GCC ACT ACG CGC CAC CTG CGG CCC GAC ATG Glu Glu Arg Leu Arg Phe Ala Thr Thr Arg His Leu Arg Pro Asp Met 795 800 805	2570
AAG ACG CAC CTG GAC CGG CCG CTG GTG GTG GAG CTG GGC CGC GAC GGC Lys Thr His Leu Asp Arg Pro Leu Val Val Glu Leu Gly Arg Asp Gly 810 815 820 825	2618
GCG CGG GGG CCC GTG GGA GGC AAA GCC CGA CCT GAG GCT GCG GAG CCC Ala Arg Gly Pro Val Gly Gly Lys Ala Arg Pro Glu Ala Ala Glu Ala 830 835 840	2666
CCC GAG GGC GTC GAC CCT CCG CGC AGG CAC CAC CGG CAC CGC GAC AAG Pro Glu Gly Val Asp Pro Pro Arg Arg His His Arg His Arg Asp Lys 845 850 855	2714
GAC AAG ACC CCC GCG GCG GGG GAC CAG GAC CGA GCA GAG GCC CCG AAG Asp Lys Thr Pro Ala Ala Gly Asp Gln Asp Arg Ala Glu Ala Pro Lys 860 865 870	2762
GCG GAG AGC GGG GAG CCC GGT GCC CGG GAG GAG CGG CCG CGG CCG CAC Ala Glu Ser Gly Glu Pro Gly Ala Arg Glu Glu Arg Pro Arg Pro His 875 880 885	2810
CGC AGC CAC AGC AAG GAG GCC GCG GGG CCC CCG GAG GCG CGG AGC GAG Arg Ser His Ser Lys Glu Ala Ala Gly Pro Pro Glu Ala Arg Ser Glu 890 895 900 905	2858
CGC GGC CGA GGC CCA GGC CCC GAG GGC GGC CGG CGG CAC CAC CGG CGC Arg Gly Arg Gly Pro Gly Pro Glu Gly Gly Arg Arg His His Arg Arg 910 915 920	2906
GGC TCC CCG GAG GAG GCG GCC GAG CGG GAG CCC CGA CGC CAC CGC GCG Gly Ser Pro Glu Glu Ala Ala Glu Arg Glu Pro Arg Arg His Arg Ala 925 930 935	2954
CAC CGG CAC CAG GAT CCG AGC AAG GAG TGC GCC GGC GCC AAG GGC GAG His Arg His Gln Asp Pro Ser Lys Glu Cys Ala Gly Ala Lys Gly Glu 940 945 950	3002
CGG CGC GCG CGG CAC CGC GGC GGC CCC CGA GCG GGG CCC CGG GAG GCG Arg Arg Ala Arg His Arg Gly Gly Pro Arg Ala Gly Pro Arg Glu Ala 955 960 965	3050
GAG AGC GGG GAG GAG CCG GCG CGG CGG CAC CGG GCC CGG CAC AAG GCG Glu Ser Gly Glu Glu Pro Ala Arg Arg His Arg Ala Arg His Lys Ala 970 975 980 985	3098
CAG CCT GCT CAC GAG GCT GTG GAG AAG GAG ACC ACG GAG AAG GAG GCC Gln Pro Ala His Glu Ala Val Glu Lys Glu Thr Thr Glu Lys Glu Ala 990 995 1000	3146

-115-

ACG GAG AAG GAG GCT GAG ATA GTG GAA GCC GAC AAG GAA AAG GAG CTC Thr Glu Lys Glu Ala Ile Val Glu Ala Asp Lys Glu Lys Glu Leu 1005 1010 1015	3194
CGG AAC CAC CAG CCC CGG GAG CCA CAC TGT GAC CTG GAG ACC AGT GGG Arg Asn His Gln Pro Arg Glu Pro His Cys Asp Leu Glu Thr Ser Gly 1020 1025 1030	3242
ACT GTG ACT GTG GGT CCC ATG CAC ACA CTG CCC AGC ACC TGT CTC CAG Thr Val Thr Val Gly Pro Met His Thr Leu Pro Ser Thr Cys Leu Gln 1035 1040 1045	3290
AAG GTG GAG GAA CAG CCA GAG GAT GCA GAC AAT CAG CGG AAC GTC ACT Lys Val Glu Glu Gln Pro Glu Asp Ala Asp Asn Gln Arg Asn Val Thr 1050 1055 1060 1065	3338
CGC ATG GGC AGT CGG CCC CCA GAC CGG AAC ACT ATT GTA CAT ATC CCA Arg Met Gly Ser Gln Pro Pro Asp Pro Asn Thr Ile Val His Ile Pro 1070 1075 1080	3386
GTG ATG CTG ACG GGC CCT CTT GGG GAA GCC ACG GTC GTT CCC AGT GGT Val Met Leu Thr Gly Pro Leu Gly Glu Ala Thr Val Val Pro Ser Gly 1085 1090 1095	3434
AAC GTG GAC CTG GAA AGC CAA GCA GAG GGG AAG AAG GAG GTG GAA GCG Asn Val Asp Leu Glu Ser Gln Ala Glu Gly Lys Lys Glu Val Glu Ala 1100 1105 1110	3482
GAT GAC GTG ATG AGG AGC GGC CCC CGG CCT ATC GTC CCA TAC AGC TCC Asp Asp Val Met Arg Ser Gly Pro Arg Pro Ile Val Pro Tyr Ser Ser 1115 1120 1125	3530
ATG TTC TGT TTA AGC CCC ACC AAC CTG CTC CGC CGC TTC TGC CAC TAC Met Phe Cys Leu Ser Pro Thr Asn Leu Leu Arg Arg Phe Cys His Tyr 1130 1135 1140 1145	3578
ATC GTG ACC ATG AGG TAC TTC GAG GTG GTC ATT CTC GTG GTC ATC GCC Ile Val Thr Met Arg Tyr Phe Glu Val Val Ile Leu Val Val Ile Ala 1150 1155 1160	3626
TTG AGC AGC ATC GCC CTG GCT GCT GAG GAC CCA GTG CGC ACA GAC TCG Leu Ser Ser Ile Ala Leu Ala Ala Glu Asp Pro Val Arg Thr Asp Ser 1165 1170 1175	3674
CCC AGG AAC AAC GCT CTG AAA TAC CTG GAT TAC ATT TTC ACT GGT GTC Pro Arg Asn Asn Ala Leu Lys Tyr Leu Asp Tyr Ile Phe Thr Gly Val 1180 1185 1190	3722
TTT ACC TTT GAG ATG GTG ATA AAG ATG ATC GAC TTG GGA CTG CTG CTT Phe Thr Phe Glu Met Val Ile Lys Met Ile Asp Leu Gly Leu Leu Leu 1195 1200 1205	3770
CAC CCT GGA GCC TAT TTC CCG GAC TTG TGG AAC ATT CTG GAC TTC ATT His Pro Gly Ala Tyr Phe Arg Asp Leu Trp Asn Ile Leu Asp Phe Ile 1210 1215 1220 1225	3818
GTG GTC AGT GGC GCC CTG GTG GCG TTT GCT TTC TCA GGA TCC AAA GGG Val Val Ser Gly Ala Leu Val Ala Phe Ala Phe Ser Gly Ser Lys Gly 1230 1235 1240	3866
AAA GAC ATC AAT ACC ATC AAG TCT CTG AGA GTC CTT CGT GTC CTG CGG Lys Asp Ile Asn Thr Ile Lys Ser Leu Arg Val Leu Arg Val Leu Arg 1245 1250 1255	3914
CCC CTC AAG ACC ATC AAA CGG CTG CCC AAG CTC AAG GCT GTG TTT GAC Pro Leu Lys Thr Ile Lys Arg Leu Pro Lys Leu Lys Ala Val Phe Asp 1260 1265 1270	3962

-116-

TGT GTG GTG AAC TCC CTG AAG AAT GTC CTC AAC ATC TTG ATT GTC TAC Cys Val Val Asn Ser Leu Lys Asn Val Leu Asn Ile Leu Ile Val Tyr 1275 1280 1285	4010
ATG CTC TTC ATG TTC ATA TTT GCC GTC ATT GCG GTG CAG CTC TTC AAA Met Leu Phe Met Phe Ile Phe Ala Val Ile Ala Val Gln Leu Phe Lys 1290 1295 1300 1305	4058
GGG AAG TTT TTC TAC TGC ACA GAT GAA TCC AAG GAG CTG GAG AGG GAC Gly Lys Phe Phe Tyr Cys Thr Asp Glu Ser Lys Glu Leu Glu Arg Asp 1310 1315 1320	4106
TGC AGG GGT CAG TAT TTG GAT TAT GAG AAG GAG GAA GTG GAA GCT CAG Cys Arg Gly Gln Tyr Leu Asp Tyr Glu Lys Glu Glu Val Glu Ala Gln 1325 1330 1335	4154
CCC AGG CAG TGG AAG AAA TAC GAC TTT CAC TAC GAC AAT GTG CTC TGG Pro Arg Gln Trp Lys Lys Tyr Asp Phe His Tyr Asp Asn Val Leu Trp 1340 1345 1350	4202
GCT CTG CTG ACG CTG TTC ACA GTG TCC ACG GGA GAA GGC TGG CCC ATG Ala Leu Leu Thr Leu Phe Thr Val Ser Thr Gly Glu Gly Trp Pro Met 1355 1360 1365	4250
GTG CTG AAA CAC TCC GTG GAT GCC ACC TAT GAG GAG CAG GGT CCA AGC Val Leu Lys His Ser Val Asp Ala Thr Tyr Glu Glu Gln Gly Pro Ser 1370 1375 1380 1385	4298
CCT GGG TAC CGC ATG GAG CTG TCC ATC TTC TAC GTG GTC TAC TTT GTG Pro Gly Tyr Arg Met Glu Leu Ser Ile Phe Tyr Val Val Tyr Phe Val 1390 1395 1400	4346
GTC TTT CCC TTC TTC GTC AAC ATC TTT GTG GCT TTG ATC ATC ATC Val Phe Pro Phe Phe Val Asn Ile Phe Val Ala Leu Ile Ile Ile 1405 1410 1415	4394
ACC TTC CAG GAG CAG GGG GAC AAG GTG ATG TCT GAA TGC AGC CTG GAG Thr Phe Gln Glu Gln Gly Asp Lys Val Met Ser Glu Cys Ser Leu Glu 1420 1425 1430	4442
AAG AAC GAG AGG GCT TGC ATT GAC TTC GCC ATC AGC GCC AAA CCC CTG Lys Asn Glu Arg Ala Cys Ile Asp Phe Ala Ile Ser Ala Lys Pro Leu 1435 1440 1445	4490
ACA CGG TAC ATG CCC CAA AAC CGG CAG TCG TTC CAG TAT AAG ACG TGG Thr Arg Tyr Met Pro Gln Asn Arg Gln Ser Phe Gln Tyr Lys Thr Trp 1450 1455 1460 1465	4538
ACA TTT GTG GTC TCC CCG CCC TTT GAA TAC TTC ATC ATG GCC ATG ATA Thr Phe Val Val Ser Pro Pro Phe Glu Tyr Phe Ile Met Ala Met Ile 1470 1475 1480	4586
GCC CTC AAC ACT GTG GTG CTG ATG ATG AAG TTC TAT GAT GCA CCC TAT Ala Leu Asn Thr Val Val Leu Met Met Lys Phe Tyr Asp Ala Pro Tyr 1485 1490 1495	4634
GAG TAC GAG CTG ATG CTG AAA TGC CTG AAC ATC GTG TTC ACA TCC ATG Glu Tyr Glu Leu Met Leu Lys Cys Leu Asn Ile Val Phe Thr Ser Met 1500 1505 1510	4682
TTC TCC ATG GAA TGC GTG CTG AAG ATC ATC GCC TTT GGG GTG CTG AAC Phe Ser Met Glu Cys Val Leu Lys Ile Ile Ala Phe Gly Val Leu Asn 1515 1520 1525	4730
TAT TTC AGA GAT GCC TGG AAT GTC TTT GAC TTT GTC ACT GTG TTG GGA Tyr Phe Arg Asp Ala Trp Asn Val Phe Asp Phe Val Thr Val Leu Gly 1530 1535 1540 1545	4778

-117-

AGT ATT ACT GAT ATT TTA GTA ACA GAG ATT GCG GAA ACG AAC AAT TTC Ser Ile Thr Asp Ile Leu Val Thr Glu Ile Ala Glu Thr Asn Asn Phe 1550 1555 1560	4826
ATC AAC CTC AGC TTC CTC CGC CTC TTT CGA GCT GCG CGG CTG ATC AAG Ile Asn Leu Ser Phe Leu Arg Leu Phe Arg Ala Ala Arg Leu Ile Lys 1565 1570 1575	4874
CTG CTC CGC CAG GGC TAC ACC ATC CGC ATC CTG CTG TGG ACC TTT GTC Leu Leu Arg Gln Gly Tyr Thr Ile Arg Ile Leu Leu Trp Thr Phe Val 1580 1585 1590	4922
CAG TCC TTC AAG GCC CTG CCC TAC GTG TGT CTG CTC ATT GCC ATG CTG Gln Ser Phe Lys Ala Leu Pro Tyr Val Cys Leu Leu Ile Ala Met Leu 1595 1600 1605	4970
TTC TTC ATC TAC GCC ATC ATC GGC ATG CAG GTG TTT GGG AAT ATT GCC Phe Phe Ile Tyr Ala Ile Ile Gly Met Gln Val Phe Gly Asn Ile Ala 1610 1615 1620 1625	5018
CTG GAT GAT GAC ACC AGC ATC AAC CGC CAC AAC AAC TTC CGG ACG TTT Leu Asp Asp Asp Thr Ser Ile Asn Arg His Asn Asn Phe Arg Thr Phe 1630 1635 1640	5066
TTG CAA GCC CTG ATG CTG CTG TTC AGG AGC GCC ACG GGG GAG GCC TGG Leu Gln Ala Leu Met Leu Leu Phe Arg Ser Ala Thr Gly Glu Ala Trp 1645 1650 1655	5114
CAC GAG ATC ATG CTG TCC TGC CTG AGC AAC CAG GCC TGT GAT GAG CAG His Glu Ile Met Leu Ser Cys Leu Ser Asn Gln Ala Cys Asp Glu Gln 1660 1665 1670	5162
GCC AAT GCC ACC GAG TGT GGA AGT GAC TTT GCC TAC TTC TAC TTC GTC Ala Asn Ala Thr Glu Cys Gly Ser Asp Phe Ala Tyr Phe Tyr Phe Val 1675 1680 1685	5210
TCC TTC ATC TTC CTG TGC TCC TTT CTG ATG TTG AAC CTC TTT GTG GCT Ser Phe Ile Phe Leu Cys Ser Phe Leu Met Leu Asn Leu Phe Val Ala 1690 1695 1700 1705	5258
GTG ATC ATG GAC AAT TTT GAG TAC CTC ACG CGG GAC TCT TCC ATC CTA Val Ile Met Asp Asn Phe Glu Tyr Leu Thr Arg Asp Ser Ser Ile Leu 1710 1715 1720	5306
GGT CCT CAC CAC TTG GAT GAG TTC ATC CGG GTC TGG GCT GAA TAC GAC Gly Pro His His Leu Asp Glu Phe Ile Arg Val Trp Ala Glu Tyr Asp 1725 1730 1735	5354
CCG GCT GCG TGT GGG CGC ATC AGT TAC AAT GAC ATG TTT GAG ATG CTG Pro Ala Ala Cys Gly Arg Ile Ser Tyr Asn Asp Met Phe Glu Met Leu 1740 1745 1750	5402
AAA CAC ATG TCC CCG CCT CTG GGG CTG GGG AAG AAA TGC CCT GCT CGA Lys His Met Ser Pro Pro Leu Gly Leu Gly Lys Lys Cys Pro Ala Arg 1755 1760 1765	5450
GTT GCT TAC AAG CGC CTG GTT CGC ATG AAC ATG CCC ATC TCC AAC GAG Val Ala Tyr Lys Arg Leu Val Arg Met Asn Met Pro Ile Ser Asn Glu 1770 1775 1780 1785	5498
GAC ATG ACT GTT CAC TTC ACG TCC ACG CTG ATG GCC CTC ATC CGG ACG Asp Met Thr Val His Phe Thr Ser Thr Leu Met Ala Leu Ile Arg Thr 1790 1795 1800	5546
GCA CTG GAG ATC AAG CTG GCC CCA GCT GGG ACA AAG CAG CAT CAG TGT Ala Leu Glu Ile Lys Leu Ala Pro Ala Gly Thr Lys Gln His Gln Cys 1805 1810 1815	5594

-118-

GAC GCG GAG TTG AGG AAG GAG ATT TCC GTT GTG TGG GCC AAT CTG CCC Asp Ala Glu Leu Arg Lys Glu Ile Ser Val Val Trp Ala Asn Leu Pro 1820 1825 1830	5642
CAG AAG ACT TTG GAC TTG CTG GTA CCA CCC CAT AAG CCT GAT GAG ATG Gln Lys Thr Leu Asp Leu Leu Val Pro Pro His Lys Pro Asp Glu Met 1835 1840 1845	5690
ACA GTG GGG AAG GTT TAT GCA GCT CTG ATG ATA TTT GAC TTC TAC AAG Thr Val Gly Lys Val Tyr Ala Ala Leu Met Ile Phe Asp Phe Tyr Lys 1850 1855 1860 1865	5738
CAG AAC AAA ACC ACC AGA GAC CAG ATG CAG CAG GCT CCT GGA GGC CTC Gln Asn Lys Thr Thr Arg Asp Gln Met Gln Gln Ala Pro Gly Gly Leu 1870 1875 1880	5786
TCC CAG ATG GGT CCT GTG TCC CTG TTC CAC CCT CTG AAG GCC ACC CTG Ser Gln Met Gly Pro Val Ser Leu Phe His Pro Leu Lys Ala Thr Leu 1885 1890 1895	5834
GAG CAG ACA CAG CCG GCT GTG CTC CGA GGA GCC CGG GTT TTC CTT CGA Glu Gln Thr Gln Pro Ala Val Leu Arg Gly Ala Arg Val Phe Leu Arg 1900 1905 1910	5882
CAG AAG AGT TCC ACC TCC CTC AGC AAT GGC GGG GCC ATA CAA AAC CAA Gln Lys Ser Ser Thr Ser Leu Ser Asn Gly Gly Ala Ile Gln Asn Gln 1915 1920 1925	5930
GAG AGT GGC ATC AAA GAG TCT GTC TCC TGG GGC ACT CAA AGG ACC CAG Glu Ser Gly Ile Lys Glu Ser Val Ser Trp Gly Thr Gln Arg Thr Gln 1930 1935 1940 1945	5978
GAT GCA CCC CAT GAG GCC AGG CCA CCC CTG GAG CGT GGC CAC TCC ACA Asp Ala Pro His Glu Ala Arg Pro Pro Leu Glu Arg Gly His Ser Thr 1950 1955 1960	6026
GAG ATC CCT GTG GGG CGG TCA GGA GCA CTG GCT GTG GAC GTT CAG ATG Glu Ile Pro Val Gly Arg Ser Gly Ala Leu Ala Val Asp Val Gln Met 1965 1970 1975	6074
CAG AGC ATA ACC CGG AGG GGC CCT GAT GGG GAG CCC CAG CCT GGG CTG Gln Ser Ile Thr Arg Arg Gly Pro Asp Gly Glu Pro Gln Pro Gly Leu 1980 1985 1990	6122
GAG AGC CAG GGT CGA GCG GCC TCC ATG CCC CGC CTT GCG GCC GAG ACT Glu Ser Gln Gly Arg Ala Ala Ser Met Pro Arg Leu Ala Ala Glu Thr 1995 2000 2005	6170
CAG CCC GTC ACA GAT GCC AGC CCC ATG AAG CGC TCC ATC TCC ACG CTG Gln Pro Val Thr Asp Ala Ser Pro Met Lys Arg Ser Ile Ser Thr Leu 2010 2015 2020 2025	6218
GCC CAG CGG CCC CGT GGG ACT CAT CTT TGC AGC ACC ACC CCG GAC CGC Ala Gln Arg Pro Arg Gly Thr His Leu Cys Ser Thr Thr Pro Asp Arg 2030 2035 2040	6266
CCA CCC CCT AGC CAG GCG TCG TCG CAC CAC CAC CAC CGC TGC CAC Pro Pro Pro Ser Gln Ala Ser Ser His His His His His Arg Cys His 2045 2050 2055	6314
CGC CGC AGG GAC AGG AAG CAG AGG TCC CTG GAG AAG GGG CCC AGC CTG Arg Arg Arg Asp Arg Lys Gln Arg Ser Leu Glu Lys Gly Pro Ser Leu 2060 2065 2070	6362
TCT GCC GAT ATG GAT GGC GCA CCA AGC AGT GCT GTG GGG CCG GGG CTG Ser Ala Asp Met Asp Gly Ala Pro Ser Ser Ala Val Gly Pro Gly Leu 2075 2080 2085	6410

-119-

CCC CCG GGA GAG GGG CCT ACA GGC TGC CGG CGG GAA CGA GAG CGC CGG Pr Pro Gly Glu Gly Pro Thr Gly Cys Arg Arg Glu Arg Glu Arg Arg 2090 2095 2100 2105	6458
CAG GAG CGG GGC CGG TCC CAG GAG CGG AGG CAG CCC TCA TCC TCC TCC Gln Glu Arg Gly Arg Ser Gln Glu Arg Arg Gln Pro Ser Ser Ser Ser 2110 2115 2120	6506
TCG GAG AAG CAG CGC TTC TAC TCC TGC GAC CGC TTT GGG GGC CGT GAG Ser Glu Lys Gln Arg Phe Tyr Ser Cys Asp Arg Phe Gly Gly Arg Glu 2125 2130 2135	6554
CCC CCG AAG CCC AAG CCC TCC CTC AGC AGC CAC CCA ACG TCG CCA ACA Pro Pro Lys Pro Lys Pro Ser Leu Ser Ser His Pro Thr Ser Pro Thr 2140 2145 2150	6602
GCT GGC CAG GAG CCG GGA CCC CAC CCA CAG GGC AGT GGT TCC GTG AAT Ala Gly Gln Glu Pro Gly Pro His Pro Gln Gly Ser Gly Ser Val Asn 2155 2160 2165	6650
GGG AGC CCC TTG CTG TCA ACA TCT GGT GCT AGC ACC CCC GGC CGC CGT Gly Ser Pro Leu Leu Ser Thr Ser Gly Ala Ser Thr Pro Gly Arg Gly 2170 2175 2180 2185	6698
GGG CGG AGG CAG CTC CCC CAG ACG CCC CTG ACT CCC CGC CCC AGC ATC Gly Arg Arg Gln Leu Pro Gln Thr Pro Leu Thr Pro Arg Pro Ser Ile 2190 2195 2200	6746
ACC TAC AAG ACG GCC AAC TCC TCA CCC ATC CAC TTC GCC GGG GCT CAG Thr Tyr Lys Thr Ala Asn Ser Ser Pro Ile His Phe Ala Gly Ala Gln 2205 2210 2215	6794
ACC AGC CTC CCT GCC TTC TCC CCA GGC CGG CTC AGC CGT GGG CTT TCC Thr Ser Leu Pro Ala Phe Ser Pro Gly Arg Leu Ser Arg Gly Leu Ser 2220 2225 2230	6842
GAA CAC AAC GCC CTG CTG CAG AGA GAC CCC CTC AGC CAG CCC CTG GCC Glu His Asn Ala Leu Leu Gln Arg Asp Pro Leu Ser Gln Pro Leu Ala 2235 2240 2245	6890
CCT GGC TCT CGA ATT GGC TCT GAC CCT TAC CTG GGG CAG CGT CTG GAC Pro Gly Ser Arg Ile Gly Ser Asp Pro Tyr Leu Gly Gln Arg Leu Asp 2250 2255 2260 2265	6938
AGT GAG GCC TCT GTC CAC GCC CTG CCT GAG GAC ACG CTC ACT TTC GAG Ser Glu Ala Ser Val His Ala Leu Pro Glu Asp Thr Leu Thr Phe Glu 2270 2275 2280	6986
GAG GCT GTG GCC ACC AAC TCG GGC CGC TCC AGG ACT TCC TAC GTG Glu Ala Val Ala Thr Asn Ser Gly Arg Ser Ser Arg Thr Ser Tyr Val 2285 2290 2295	7034
TCC TCC CTG ACC TCC CAG TCT CAC CCT CTC CGC CGC GTG CCC AAC GGT Ser Ser Leu Thr Ser Gln Ser His Pro Leu Arg Arg Val Pro Asn Gly 2300 2305 2310	7082
TAC CAC TGC ACC CTG GGA CTC AGC TCG GGT GGC CGA GCA CGG CAC AGC Tyr His Cys Thr Leu Gly Leu Ser Ser Gly Gly Arg Ala Arg His Ser 2315 2320 2325	7130
TAC CAC CAC CCT GAC CAA GAC CAC TGG TGC TAGCTGCACC GTGACCGCTC Tyr His His Pro Asp Gln Asp His Trp Cys 2330 2335 234	7180
AGACGCCCTGC ATGCAGCAGG CGTGTGTTCC AGTGGATGAG TTTTATCATC CACACGGGGC AGTCGGCCCT CGGGGGAGGC CTTGCCACC TTGGTGAGGC TCCTGTGGCC CCTCCCTCCCC	7240
	7300

-120-

CCTCCTCCCC TCTTTTACTC TAGACGACGA ATAAAGCCCT GTTGCTTGAG TGTACGTACC	7360
GC	7362

## (2) INFORMATION FOR SEQ ID NO:8:

## (i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 7175 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: double
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (genomic)

## (ix) FEATURE:

- (A) NAME/KEY: CDS
- (B) LOCATION: 144..6857

## (ix) FEATURE:

- (A) NAME/KEY: 5'UTR
- (B) LOCATION: 1..143

## (ix) FEATURE:

- (A) NAME/KEY: 3'UTR
- (B) LOCATION: 6855..7175

## (xi) SEQUENCE DESCRIPTION: SEQ ID NO:8:

GGGGCGGCCGG CTGGCGCGGT GGGGCCGGGC GAGGTCCGTG CGGTCCCGGC GGCTCCGTGG	60
--	----

CTGCTCCGCT CTGAGCGCCT GCGCGCCCCG CGCCCTCCCT GCCGGGGCCG CTGGGCGCGG	120
---	-----

GATGCACGCG GGGCCCGGGA GCC ATG GTC CGC TTC GGG GAC GAG CTG GGC Met Val Arg Phe Gly Asp Glu Leu Gly	170
--	-----

1

5

GGC CGC TAT GGA GGC CCC GGC GGC GGA GAG CGG GCC CGG GGC GGC GGG Gly Arg Tyr Gly Gly Pro Gly Gly Glu Arg Ala Arg Gly Gly Gly	218
--	-----

10 15 20 25
-------------

GCC GGC GGG GCG GGG GGC CCG GGT CCC GGG GGG CTG CAG CCC GGC CAG Ala Gly Ala Gly Gly Pro Gly Pro Gly Gly Leu Gln Pro Gly Gln	266
--	-----

30

35

40

CGG GTC CTC TAC AAG CAA TCG ATC GCG CAG CGC GCG CGG ACC ATG GCG Arg Val Leu Tyr Lys Gln Ser Ile Ala Gln Arg Ala Arg Thr Met Ala	314
--	-----

45

50

55

CTG TAC AAC CCC ATC CCG GTC AAG CAG AAC TGC TTC ACC GTC AAC CGC Leu Tyr Asn Pro Ile Pro Val Lys Gln Asn Cys Phe Thr Val Asn Arg	362
--	-----

60

65

70

TCG CTC TTC GTC TTC AGC GAG GAC AAC GTC GTC CGC AAA TAC GCG AAG Ser Leu Phe Val Phe Ser Glu Asp Asn Val Val Arg Lys Tyr Ala Lys	410
--	-----

75

80

85

CGC ATC ACC GAG TGG CCT CCA TTC GAG AAT ATG ATC CTG GCC ACC ATC Arg Ile Thr Glu Trp Pro Pro Phe Glu Asn Met Ile Leu Ala Thr Ile	458
--	-----

90

95

100

105

ATC GCC AAC TGC ATC GTG CTG GCC CTG GAG CAG CAC CTC CCT GAT GGG Ile Ala Asn Cys Ile Val Leu Ala Leu Glu Gln His Leu Pro Asp Gly	506
--	-----

110

115

120

GAC AAA ACG CCC ATG TCC GAG CGG CTG GAC GAC ACG GAG CCC TAT TTC Asp Lys Thr Pro Met Ser Glu Arg Leu Asp Asp Thr Glu Pro Tyr Phe	554
--	-----

125

130

135

-121-

ATC GGG ATC TTT TGC TTC GAG GCA GGG ATC AAA ATC ATC GCT CTG GGC Ile Gly Ile Phe Cys Phe Glu Ala Gly Ile Lys Ile Ile Ala Leu Gly 140 145 150	602
TTT GTC TTC CAC AAG GGC TCT TAC CTG CGG AAC GGC TGG AAC GTC ATG Phe Val Phe His Lys Gly Ser Tyr Leu Arg Asn Gly Trp Asn Val Met 155 160 165	650
GAC TTC GTG GTC CTC ACA GGG ATC CTT GCC ACG GCT GGA ACT GAC Asp Phe Val Val Val Leu Thr Gly Ile Leu Ala Thr Ala Gly Thr Asp 170 175 180 185	698
TTC GAC CTG CGA ACA CTG AGG GCT GTG CGT GTG CTG AGG CCC CTG AAG Phe Asp Leu Arg Thr Leu Arg Ala Val Arg Val Leu Arg Pro Leu Lys 190 195 200	746
CTG GTG TCT GGG ATT CCA AGT TTG CAG GTG GTG CTC AAG TCC ATC ATG Leu Val Ser Gly Ile Pro Ser Leu Gln Val Val Leu Lys Ser Ile Met 205 210 215	794
AAG GCC ATG GTT CCA CTC CTG CAG ATT GGG CTG CTT CTC TTC TTT GCC Lys Ala Met Val Pro Leu Leu Gln Ile Gly Leu Leu Leu Phe Phe Ala 220 225 230	842
ATC CTC ATG TTT GCC ATC ATT GGC CTG GAG TTC TAC ATG GGC AAG TTC Ile Leu Met Phe Ala Ile Ile Gly Leu Glu Phe Tyr Met Gly Lys Phe 235 240 245	890
CAC AAG GCC TGT TTC CCC AAC AGC ACA GAT GCG GAG CCC GTG GGT GAC His Lys Ala Cys Phe Pro Asn Ser Thr Asp Ala Glu Pro Val Gly Asp 250 255 260 265	938
TTC CCC TGT GGC AAG GAG GCC CCA GCC CGG CTG TGC GAG GGC GAC ACT Phe Pro Cys Gly Lys Glu Ala Pro Ala Arg Leu Cys Glu Gly Asp Thr 270 275 280	986
GAG TGC CGG GAG TAC TGG CCA GGA CCC AAC TTT GGC ATC ACC AAC TTT Glu Cys Arg Glu Tyr Trp Pro Gly Pro Asn Phe Gly Ile Thr Asn Phe 285 290 295	1034
GAC AAT ATC CTG TTT GCC ATC TTG ACG GTG TTC CAG TGC ATC ACC AAC ATG Asp Asn Ile Leu Phe Ala Ile Leu Thr Val Phe Gln Cys Ile Thr Met 300 305 310	1082
GAG GGC TGG ACT GAC ATC CTC TAT AAT ACA AAC GAT GCG GCC GGC AAC Glu Gly Trp Thr Asp Ile Leu Tyr Asn Thr Asn Asp Ala Ala Gly Asn 315 320 325	1130
ACC TGG AAC TGG CTC TAC TTC ATC CCT CTC ATC ATC ATC GGC TCC TTC Thr Trp Asn Trp Leu Tyr Phe Ile Pro Leu Ile Ile Ile Gly Ser Phe 330 335 340 345	1178
TTC ATG CTC AAC CTG GTG CTG GGC GTG CTC TCG GGG GAG TTT GCC AAG Phe Met Leu Asn Leu Val Leu Gly Val Leu Ser Gly Glu Phe Ala Lys 350 355 360	1226
GAG CGA GAG AGG GTG GAG AAC CGC CGC GCC TTC CTG AAG CTG CGC CGG Glu Arg Glu Arg Val Glu Asn Arg Arg Ala Phe Leu Lys Leu Arg Arg 365 370 375	1274
CAG CAG CAG ATC GAG CGA GAG CTC AAC GGG TAC CTG GAG TGG ATC TTC Gln Gln Ile Glu Arg Glu Leu Asn Gly Tyr Leu Glu Trp Ile Phe 380 385 390	1322
AAG GCG GAG GAA GTC ATG CTG GCC GAG GAG GAC AGG AAT GCA GAG GAG Lys Ala Glu Glu Val Met Leu Ala Glu Glu Asp Arg Asn Ala Glu Glu 395 400 405	1370

-122-

AAG TCC CCT TTG GAC GTG CTG AAG AGA GCG GCC ACC AAG AAG AGC AGA Lys Ser Pro Leu Asp Val Leu Lys Arg Ala Ala Thr Lys Lys Ser Arg 410 415 420 425	1418
AAT GAC CTG ATC CAC GCA GAG GAG GGA GAG GAC CGG TTT GCA GAT CTC Asn Asp Leu Ile His Ala Glu Glu Gly Glu Asp Arg Phe Ala Asp Leu 430 435 440	1466
TGT GCT GTT GGA TCC CCC TTC GCC CGC GCC AGC CTC AAG AGC GGG AAG Cys Ala Val Gly Ser Pro Phe Ala Arg Ala Ser Leu Lys Ser Gly Lys 445 450 455	1514
ACA GAG AGC TCG TCA TAC TTC CGG AGG AAG GAG AAG ATG TTC CGG TTT Thr Glu Ser Ser Tyr Phe Arg Arg Lys Glu Lys Met Phe Arg Phe 460 465 470	1562
TTT ATC CGG CGC ATG GTG AAG GCT CAG AGC TTC TAC TGG GTG GTG CTG Phe Ile Arg Arg Met Val Lys Ala Gln Ser Phe Tyr Trp Val Val Leu 475 480 485	1610
TGC GTG GTG GCC CTG AAC ACA CTG TGT GTG GCC ATG GTG CAT TAC AAC Cys Val Val Ala Leu Asn Thr Leu Cys Val Ala Met Val His Tyr Asn 490 495 500 505	1658
CAG CCG CGG CGG CTT ACC ACG ACC CTG TAT TTT GCA GAG TTT GTT TTC Gln Pro Arg Arg Leu Thr Thr Leu Tyr Phe Ala Glu Phe Val Phe 510 515 520	1706
CTG GGT CTC TTC CTC ACA GAG ATG TCC CTG AAG ATG TAT GGC CTG GGG Leu Gly Leu Phe Leu Thr Glu Met Ser Leu Lys Met Tyr Gly Leu Gly 525 530 535	1754
CCC AGA AGC TAC TTC CGG TCC TCC AAC TGC TTC GAC TTT GGG GTC Pro Arg Ser Tyr Phe Arg Ser Ser Phe Asn Cys Phe Asp Phe Gly Val 540 545 550	1802
ATC GTG GGG AGC GTC TTT GAA GTG GTC TGG GCG GCC ATC AAG CCG GGA Ile Val Gly Ser Val Phe Glu Val Val Trp Ala Ala Ile Lys Pro Gly 555 560 565	1850
AGC TCC TTT GGG ATC AGT GTG CTG CGG GCC CTC CGC CTG CTG AGG ATC Ser Ser Phe Gly Ile Ser Val Leu Arg Ala Leu Arg Leu Leu Arg Ile 570 575 580 585	1898
TTC AAA GTC ACG AAG TAC TGG AGC TCC CTG CGG AAC CTG GTG GTG TCC Phe Lys Val Thr Lys Tyr Trp Ser Ser Leu Arg Asn Leu Val Val Ser 590 595 600	1946
CTG CTG AAC TCC ATG AAG TCC ATC ATC AGC CTG CTC TTC TTG CTC TTC Leu Leu Asn Ser Met Lys Ser Ile Ile Ser Leu Leu Phe Leu Leu Phe 605 610 615	1994
CTG TTC ATT GTG GTC TTC GCC CTG CTG GGG ATG CAG CTG TTT GGG GGA Leu Phe Ile Val Val Phe Ala Leu Leu Gly Met Gln Leu Phe Gly Gly 620 625 630	2042
CAG TTC AAC TTC CAG GAT GAG ACT CCC ACA ACC AAC TTC GAC ACC TTC Gln Phe Asn Phe Gln Asp Glu Thr Pro Thr Thr Asn Phe Asp Thr Phe 635 640 645	2090
CCT GCC GCC ATC CTC ACT GTC TTC CAG ATC CTG ACG GGA GAG GAC TGG Pro Ala Ala Ile Leu Thr Val Phe Gln Ile Leu Thr Gly Glu Asp Trp 650 655 660 665	2138
AAT GCA GTG ATG TAT CAC GGG ATC GAA TCG CAA GGC GGC GTC AGC AAA Asn Ala Val Met Tyr His Gly Ile Glu Ser Gln Gly Gly Val Ser Lys 670 675 680	2186

-123-

GCC ATG TTC TCG TCC TTT TAC TTC ATT GTC CTG ACA CTG TTC GGA AAC Gly Met Phe Ser Ser Phe Tyr Phe Ile Val Leu Thr Leu Phe Gly Asn 685 690 695	2234
TAC ACT CTG CTG AAT GTC TTT CTG GCC ATC GCT GTG GAC AAC CTG GCC Tyr Thr Leu Leu Asn Val Phe Leu Ala Ile Ala Val Asp Asn Leu Ala 700 705 710	2282
AAC GCC CAA GAG CTG ACC AAG GAT GAA GAG GAG ATG GAA GAA GCA GCC Asn Ala Gln Glu Leu Thr Lys Asp Glu Glu Met Glu Glu Ala Ala 715 720 725	2330
AAT CAG AAG CTT GCT CTG CAA AAG GCC AAA GAA GTG GCT GAA GTC AGC Asn Gln Lys Leu Ala Leu Gln Lys Ala Lys Glu Val Ala Glu Val Ser 730 735 740 745	2378
CCC ATG TCT GCC GCG AAC ATC TCC ATC GCC AGG CAG CAG AAC TCG Pro Met Ser Ala Ala Asn Ile Ser Ile Ala Ala Arg Gln Gln Asn Ser 750 755 760	2426
GCC AAG GCG CGC TCG GTG TGG GAG CAG CGG GCC AGC CAG CTA CGG CTG Ala Lys Ala Arg Ser Val Trp Glu Gln Arg Ala Ser Gln Leu Arg Leu 765 770 775	2474
CAG AAC CTG CGG GCC AGC TGC GAG GCG CTG TAC AGC GAG ATG GAC CCC Gln Asn Leu Arg Ala Ser Cys Glu Ala Leu Tyr Ser Glu Met Asp Pro 780 785 790	2522
GAG GAG CGG CTG CGC TTC GCC ACT ACG CGC CAC CTG CGG CCC GAC ATG Glu Glu Arg Leu Arg Phe Ala Thr Thr Arg His Leu Arg Pro Asp Met 795 800 805	2570
AAG ACG CAC CTG GAC CGG CCG CTG GTG GTG GAG CTG GGC CGC GAC GGC Lys Thr His Leu Asp Arg Pro Leu Val Val Glu Leu Gly Arg Asp Gly 810 815 820 825	2618
GCG CGG GGG CCC GTG GGA GGC AAA GCC CGA CCT GAG GCT GCG GAG GCC Ala Arg Gly Pro Val Gly Gly Lys Ala Arg Pro Glu Ala Ala Glu Ala 830 835 840	2666
CCC GAG GGC GTC GAC CCT CCG CGC AGG CAC CAC CGG CAC CGC GAC AAG Pro Glu Gly Val Asp Pro Pro Arg Arg His His Arg His Arg Asp Lys 845 850 855	2714
GAC AAG ACC CCC GCG GCG GGG GAC CAG GAC CGA GCA GAG GCC CCG AAG Asp Lys Thr Pro Ala Ala Gly Asp Gln Asp Arg Ala Glu Ala Pro Lys 860 865 870	2762
GCG GAG AGC GGG GAG CCC GGT GCC CGG GAG GAG CGG CCG CGG CCG CAC Ala Glu Ser Gly Glu Pro Gly Ala Arg Glu Glu Arg Pro Arg Pro His 875 880 885	2810
CGC AGC CAC AGC AAG GAG GCC GCG GGG CCC CCG GAG GCG CGG AGC GAG Arg Ser His Ser Lys Glu Ala Ala Gly Pro Pro Glu Ala Arg Ser Glu 890 895 900 905	2858
CGC GGC CGA GGC CCA GGC CCC GAG GGC GGC CGG CGG CAC CAC CGG CGC Arg Gly Arg Gly Pro Gly Pro Glu Gly Gly Arg Arg His His Arg Arg 910 915 920	2906
GCC TCC CCG GAG GAG GCG GCC GAG CGG GAG CCC CGA CGC CAC CGC GCG Gly Ser Pro Glu Glu Ala Ala Glu Arg Glu Pro Arg Arg His Arg Ala 925 930 935	2954
CAC CGG CAC CAG GAT CCG AGC AAG GAG TGC GCC GGC GCC AAG GGC GAG His Arg His Gln Asp Pro Ser Lys Glu Cys Ala Gly Ala Lys Gly Glu 940 945 950	3002

-124-

CGG CGC GCG CGG CAC CGC GGC GGC CCC CGA GCG GGG CCC CGG GAG GCG Arg Arg Ala Arg His Arg Gly Gly Pro Arg Ala Gly Pro Arg Glu Ala 955 960 965	3050
GAG AGC GGG GAG GAG CCG GCG CGG CGG CAC CGG GCC CGG CAC AAG GCG Glu Ser Gly Glu Glu Pro Ala Arg Arg His Arg Ala Arg His Lys Ala 970 975 980 985	3098
CAG CCT GCT CAC GAG GCT GTG GAG AAG GAG ACC ACG GAG AAG GAG GCC Gln Pro Ala His Glu Ala Val Glu Lys Glu Thr Thr Glu Lys Glu Ala 990 995 1000	3146
ACG GAG AAG GAG GCT GAG ATA GTG GAA GCC GAC AAG GAA AAG GAG CTC Thr Glu Lys Glu Ala Glu Ile Val Glu Ala Asp Lys Glu Lys Glu Leu 1005 1010 1015	3194
CGG AAC CAC CAG CCC CGG GAG CCA CAC TGT GAC CTG GAG ACC AGT GGG Arg Asn His Gln Pro Arg Glu Pro His Cys Asp Leu Glu Thr Ser Gly 1020 1025 1030	3242
ACT GTG ACT GTG GGT CCC ATG CAC ACA CTG CCC AGC ACC TGT CTC CAG Thr Val Thr Val Gly Pro Met His Thr Leu Pro Ser Thr Cys Leu Gln 1035 1040 1045	3290
AAG GTG GAG GAA CAG CCA GAG GAT GCA GAC AAT CAG CGG AAC GTC ACT Lys Val Glu Glu Gln Pro Glu Asp Ala Asp Asn Gln Arg Asn Val Thr 1050 1055 1060 1065	3338
CGC ATG GGC AGT CAG CCC CCA GAC CCG AAC ACT ATT GTA CAT ATC CCA Arg Met Gly Ser Gln Pro Pro Asp Pro Asn Thr Ile Val His Ile Pro 1070 1075 1080	3386
GTG ATG CTG ACG GGC CCT CTT GGG GAA GCC ACG GTC GTT CCC AGT GGT Val Met Leu Thr Gly Pro Leu Gly Glu Ala Thr Val Val Pro Ser Gly 1085 1090 1095	3434
AAC GTG GAC CTG GAA AGC CAA GCA GAG GGG AAG AAG GAG GTG GAA GCG Asn Val Asp Leu Glu Ser Gln Ala Glu Gly Lys Lys Glu Val Glu Ala 1100 1105 1110	3482
GAT GAC GTG ATG AGG AGC GGC CCC CGG CCT ATC GTC CCA TAC AGC TCC Asp Asp Val Met Arg Ser Gly Pro Arg Pro Ile Val Pro Tyr Ser Ser 1115 1120 1125	3530
ATG TTC TGT TTA AGC CCC ACC AAC CTG CTC CGC CGC TTC TGC CAC TAC Met Phe Cys Leu Ser Pro Thr Asn Leu Leu Arg Arg Phe Cys His Tyr 1130 1135 1140 1145	3578
ATC GTG ACC ATG AGG TAC TTC GAG GTG GTC ATT CTC GTG GTC ATC GCC Ile Val Thr Met Arg Tyr Phe Glu Val Val Ile Leu Val Val Ile Ala 1150 1155 1160	3626
TTG AGC AGC ATC GCC CTG GCT GAG GAC CCA GTG CGC ACA GAC TCG Leu Ser Ser Ile Ala Leu Ala Ala Glu Asp Pro Val Arg Thr Asp Ser 1165 1170 1175	3674
CCC AGG AAC AAC GCT CTG AAA TAC CTG GAT TAC ATT TTC ACT GGT GTC Pro Arg Asn Asn Ala Leu Lys Tyr Leu Asp Tyr Ile Phe Thr Gly Val 1180 1185 1190	3722
TTT ACC TTT GAG ATG GTG ATA AAG ATG ATC GAC TTG GGA CTG CTG CTT Phe Thr Phe Glu Met Val Ile Lys Met Ile Asp Leu Gly Leu Leu Leu 1195 1200 1205	3770
CAC CCT GGA GCC TAT TTC CGG GAC TTG TGG AAC ATT CTG GAC TTC ATT His Pro Gly Ala Tyr Phe Arg Asp Leu Trp Asn Ile Leu Asp Phe Ile 1210 1215 1220 1225	3818

-125-

GTC GTC AGT GGC GCC CTG TG GCG TTT GCT TTC TCA GGA TCC AAA GGG Val Val Ser Gly Ala Leu Val Ala Phe Ala Phe Ser Gly Ser Lys Gly 1230 1235 1240	3866
AAA GAC ATC AAT ACC ATC AAG TCT CTG AGA GTC CTT CGT GTC CTG CGG Lys Asp Ile Asn Thr Ile Lys Ser Leu Arg Val Leu Arg Val Leu Arg 1245 1250 1255	3914
CCC CTC AAG ACC ATC AAA CGG CTG CCC AAG CTC AAG GCT GTG TTT GAC Pro Leu Lys Thr Ile Lys Arg Leu Pro Lys Leu Lys Ala Val Phe Asp 1260 1265 1270	3962
TGT GTG GTG AAC TCC CTG AAG AAT GTC CTC AAC ATC TTG ATT GTC TAC Cys Val Val Asn Ser Leu Lys Asn Val Leu Asn Ile Leu Ile Val Tyr 1275 1280 1285	4010
ATG CTC TTC ATG TTC ATA TTT GCC GTC ATT GCG GTG CAG CTC TTC AAA Met Leu Phe Met Phe Ile Phe Ala Val Ile Ala Val Gln Leu Phe Lys 1290 1295 1300 1305	4058
GGG AAG TTT TTC TAC TGC ACA GAT GAA TCC AAG GAG CTG GAG AGG GAC Gly Lys Phe Phe Tyr Cys Thr Asp Glu Ser Lys Glu Leu Glu Arg Asp 1310 1315 1320	4106
TGC AGG GGT CAG TAT TTG GAT TAT GAG AAG GAG GAA GTG GAA GCT CAG Cys Arg Gly Gln Tyr Leu Asp Tyr Glu Lys Glu Glu Val Glu Ala Gln 1325 1330 1335	4154
CCC AGG CAG TGG AAG AAA TAC GAC TTT CAC TAC GAC AAT GTG CTC TGG Pro Arg Gln Trp Lys Lys Tyr Asp Phe His Tyr Asp Asn Val Leu Trp 1340 1345 1350	4202
GCT CTG CTG ACG CTG TTC ACA GTG TCC ACG GGA GAA GGC TGG CCC ATG Ala Leu Leu Thr Leu Phe Thr Val Ser Thr Gly Glu Gly Trp Pro Met 1355 1360 1365	4250
GTG CTG AAA CAC TCC GTG GAT GCC ACC TAT GAG GAG CAG GGT CCA AGC Val Leu Lys His Ser Val Asp Ala Thr Tyr Glu Glu Gln Gly Pro Ser 1370 1375 1380 1385	4298
CCT GGG TAC CGC ATG GAG CTG TCC ATC TTC TAC GTG GTC TAC TTT GTG Pro Gly Tyr Arg Met Glu Leu Ser Ile Phe Tyr Val Val Tyr Phe Val 1390 1395 1400	4346
GTC TTT CCC TTC TTC GTC AAC ATC TTT GTG GCT TTG ATC ATC ATC Val Phe Pro Phe Phe Val Asn Ile Phe Val Ala Leu Ile Ile Ile 1405 1410 1415	4394
ACC TTC CAG GAG CAG GGG GAC AAG GTG ATG TCT GAA TGC AGC CTG GAG Thr Phe Gln Gln Gly Asp Lys Val Met Ser Glu Cys Ser Leu Glu 1420 1425 1430	4442
AAG AAC GAG AGG GCT TGC ATT GAC TTC GCC ATC AGC GCC AAA CCC CTG Lys Asn Glu Arg Ala Cys Ile Asp Phe Ala Ile Ser Ala Lys Pro Leu 1435 1440 1445	4490
ACA CGG TAC ATG CCC CAA AAC CGG CAG TCG TTC CAG TAT AAG ACG TGG Thr Arg Tyr Met Pro Gln Asn Arg Gln Ser Phe Gln Tyr Lys Thr Trp 1450 1455 1460 1465	4538
ACA TTT GTG GTC TCC CCG CCC TTT GAA TAC TTC ATC ATG GCC ATG ATA Thr Phe Val Val Ser Pro Pro Phe Glu Tyr Phe Ile Met Ala Met Ile 1470 1475 1480	4586
GCC CTC AAC ACT GTG GTG CTG ATG ATG AAG TTC TAT GAT GCA CCC TAT Ala Leu Asn Thr Val Val Leu Met Met Lys Phe Tyr Asp Ala Pro Tyr 1485 1490 1495	4634

-126-

GAG TAC GAG CTG ATG CTG AAA TGC CTG AAC ATC GTG TTC ACA TCC ATG Glu Tyr Glu Leu Met Leu Lys Cys Leu Asn Ile Val Phe Thr Ser Met 1500 1505 1510	4682
TTC TCC ATG GAA TGC GTG CTG AAG ATC ATC GCC TTT GGG GTG CTG AAC Phe Ser Met Glu Cys Val Leu Lys Ile Ile Ala Phe Gly Val Leu Asn 1515 1520 1525	4730
TAT TTC AGA GAT GCC TGG AAT GTC TTT GAC TTT GTC ACT GTG TTG GGA Tyr Phe Arg Asp Ala Trp Asn Val Phe Asp Phe Val Thr Val Leu Gly 1530 1535 1540 1545	4778
AGT ATT ACT GAT ATT TTA GTA ACA GAG ATT GCG GAA ACG AAC AAT TTC Ser Ile Thr Asp Ile Leu Val Thr Glu Ile Ala Glu Thr Asn Asn Phe 1550 1555 1560	4826
ATC AAC CTC AGC TTC CTC CGC CTC TTT CGA GCT GCG CGG CTG ATC AAG Ile Asn Leu Ser Phe Leu Arg Leu Phe Arg Ala Ala Arg Leu Ile Lys 1565 1570 1575	4874
CTG CTC CGC CAG GGC TAC ACC ATC CGC ATC CTG CTG TGG ACC TTT GTC Leu Leu Arg Gln Gly Tyr Thr Ile Arg Ile Leu Leu Trp Thr Phe Val 1580 1585 1590	4922
CAG TCC TTC AAG GCC CTG CCC TAC GTG TGT CTG CTC ATT GCC ATG CTG Gln Ser Phe Lys Ala Leu Pro Tyr Val Cys Leu Leu Ile Ala Met Leu 1595 1600 1605	4970
TTC TTC ATC TAC GCC ATC ATC GGC ATG CAG GTG TTT GGG AAT ATT GCC Phe Phe Ile Tyr Ala Ile Ile Gly Met Gln Val Phe Gly Asn Ile Ala 1610 1615 1620 1625	5018
CTG GAT GAT GAC ACC AGC ATC AAC CGC CAC AAC AAC TTC CGG ACG TTT Leu Asp Asp Asp Thr Ser Ile Asn Arg His Asn Asn Phe Arg Thr Phe 1630 1635 1640	5066
TTG CAA GCC CTG ATG CTG CTG TTC AGG AGC GCC ACG GGG GAG GCC TGG Leu Gln Ala Leu Met Leu Leu Phe Arg Ser Ala Thr Gly Glu Ala Trp 1645 1650 1655	5114
CAC GAG ATC ATG CTG TCC TGC CTG AGC AAC CAG GCC TGT GAT GAG CAG His Glu Ile Met Leu Ser Cys Leu Ser Asn Gln Ala Cys Asp Glu Gln 1660 1665 1670	5162
GCC AAT GCC ACC GAG TGT GGA AGT GAC TTT GCC TAC TTC TAC TTC GTC Ala Asn Ala Thr Glu Cys Gly Ser Asp Phe Ala Tyr Phe Tyr Phe Val 1675 1680 1685	5210
TCC TTC ATC TTC CTG TGC TCC TTT CTG ATG TTG AAC CTC TTT GTG GCT Ser Phe Ile Phe Leu Cys Ser Phe Leu Met Leu Asn Leu Phe Val Ala 1690 1695 1700 1705	5258
GTG ATC ATG GAC AAT TTT GAG TAC CTC ACG CGG GAC TCT TCC ATC CTA Val Ile Met Asp Asn Phe Glu Tyr Leu Thr Arg Asp Ser Ser Ile Leu 1710 1715 1720	5306
GGT CCT CAC CAC TTG GAT GAG TTC ATC CGG GTC TGG GCT GAA TAC GAC Gly Pro His His Leu Asp Glu Phe Ile Arg Val Trp Ala Glu Tyr Asp 1725 1730 1735	5354
CCG GCT GCG TGT GGG CGC ATC AGT TAC AAT GAC ATG TTT GAG ATG CTG Pro Ala Ala Cys Gly Arg Ile Ser Tyr Asn Asp Met Phe Glu Met Leu 1740 1745 1750	5402
AAA CAC ATG TCC CCG CCT CTG GGG CTG GGG AAG AAA TGC CCT GCT CGA Lys His Met Ser Pro Pro Leu Gly Leu Gly Lys Lys Cys Pro Ala Arg 1755 1760 1765	5450

-127-

GTT GCT TAC AAG CGC CTG GTT CGC ATG AAC ATG CCC ATC TCC AAC GAG Val Ala Tyr Lys Arg Leu Val Arg Met Asn Met Pro Ile Ser Asn Glu 1770 1775 1780 1785	5498
GAC ATG ACT GTT CAC TTC ACG TCC ACG CTG ATG GCC CTC ATC CGG ACG Asp Met Thr Val His Phe Thr Ser Thr Leu Met Ala Leu Ile Arg Thr 1790 1795 1800	5546
GCA CTG GAG ATC AAG CTG GCC CCA GCT GGG ACA AAG CAG CAT CAG TGT Ala Leu Glu Ile Lys Leu Ala Pro Ala Gly Thr Lys Gln His Gln Cys 1805 1810 1815	5594
GAC GCG GAG TTG AGG AAG GAG ATT TCC GTT GTG TGG GCC AAT CTG CCC Asp Ala Glu Leu Arg Lys Glu Ile Ser Val Val Trp Ala Asn Leu Pro 1820 1825 1830	5642
CAG AAG ACT TTG GAC TTG CTG GTA CCA CCC CAT AAG CCT GAT GAG ATG Gln Lys Thr Leu Asp Leu Leu Val Pro Pro His Lys Pro Asp Glu Met 1835 1840 1845	5690
ACA GTG GGG AAG GTT TAT GCA GCT CTG ATG ATA TTT GAC TTC TAC AAG Thr Val Gly Lys Val Tyr Ala Ala Leu Met Ile Phe Asp Phe Tyr Lys 1850 1855 1860 1865	5738
CAG AAC AAA ACC ACC AGA GAC CAG ATG CAG CAG GCT CCT GGA GGC CTC Gln Asn Lys Thr Thr Arg Asp Gln Met Gln Gln Ala Pro Gly Gly Leu 1870 1875 1880	5786
TCC CAG ATG GGT CCT GTG TCC CTG TTC CAC CCT CTG AAG GCC ACC CTG Ser Gln Met Gly Pro Val Ser Leu Phe His Pro Leu Lys Ala Thr Leu 1885 1890 1895	5834
GAG CAG ACA CAG CCG GCT GTG CTC CGA GGA GCC CGG GTT TTC CTT CGA Glu Gln Thr Gln Pro Ala Val Leu Arg Gly Ala Arg Val Phe Leu Arg 1900 1905 1910	5882
CAG AAG AGT TCC ACC TCC CTC AGC AAT GGC GGG GCC ATA CAA AAC CAA Gln Lys Ser Ser Thr Ser Leu Ser Asn Gly Gly Ile Gln Asn Gln 1915 1920 1925	5930
GAG AGT GGC ATC AAA GAG TCT GTC TCC TGG GGC ACT CAA AGG ACC CAG Glu Ser Gly Ile Lys Glu Ser Val Ser Trp Gly Thr Gln Arg Thr Gln 1930 1935 1940 1945	5978
GAT GCA CCC CAT GAG GCC AGG CCA CCC CTG GAG CGT GGC CAC TCC ACA Asp Ala Pro His Glu Ala Arg Pro Pro Leu Glu Arg Gly His Ser Thr 1950 1955 1960	6026
GAG ATC CCT GTG GGG CGG TCA GGA GCA CTG GCT GTG GAC GTT CAG ATG Glu Ile Pro Val Gly Arg Ser Gly Ala Leu Ala Val Asp Val Gln Met 1965 1970 1975	6074
CAG AGC ATA ACC CGG AGG GGC CCT GAT GGG GAG CCC CAG CCT GGG CTG Gln Ser Ile Thr Arg Arg Gly Pro Asp Gly Glu Pro Gln Pro Gly Leu 1980 1985 1990	6122
GAG AGC CAG GGT CGA GCG GCC TCC ATG CCC CGC CTT GCG GCC GAG ACT Glu Ser Gln Gly Arg Ala Ala Ser Met Pro Arg Leu Ala Ala Glu Thr 1995 2000 2005	6170
CAG CCC GTC ACA GAT GCC AGC CCC ATG AAG CGC TCC ATC TCC ACG CTG Gln Pro Val Thr Asp Ala Ser Pro Met Lys Arg Ser Ile Ser Thr Leu 2010 2015 2020 2025	6218
GCC CAG CGG CCC CGT GGG ACT CAT CTT TGC AGC ACC ACC CCG GAC CGC Ala Gln Arg Pro Arg Gly Thr His Leu Cys Ser Thr Thr Pro Asp Arg 2030 2035 2040	6266

-128-

CCA CCC CCT AGC CAG GCG TCG TCG CAC CAC CAC CAC CGC TGC CAC Pro Pro Pro Ser Gln Ala Ser Ser His His His His His Arg Cys His 2045 2050 2055	6314
CGC CGC AGG GAC AGG AAG CAG AGG TCC CTG GAG AAG GGG CCC AGC CTG Arg Arg Arg Asp Arg Lys Gln Arg Ser Leu Glu Lys Gly Pro Ser Leu 2060 2065 2070	6362
TCT GCC GAT ATG GAT GGC GCA CCA AGC AGT GCT GTG GGG CCG GGG CTG Ser Ala Asp Met Asp Gly Ala Pro Ser Ser Ala Val Gly Pro Gly Leu 2075 2080 2085	6410
CCC CGG GGA GAG GGG CCT ACA GGC TGC CGG CGG GAA CGA GAG CGC CGG Pro Pro Gly Glu Gly Pro Thr Gly Cys Arg Arg Glu Arg Glu Arg Arg 2090 2095 2100 2105	6458
CAG GAG CGG GGC CGG TCC CAG GAG CGG AGG CAG CCC TCA TCC TCC TCC Gln Glu Arg Gly Arg Ser Gln Glu Arg Arg Gln Pro Ser Ser Ser Ser 2110 2115 2120	6506
TCG GAG AAG CAG CGC TTC TAC TCC TGC GAC CGC TTT GGG GGC CGT GAG Ser Glu Lys Gln Arg Phe Tyr Ser Cys Asp Arg Phe Gly Gly Arg Glu 2125 2130 2135	6554
CCC CCG AAG CCC AAG CCC TCC CTC AGC AGC CAC CCA ACG TCG CCA ACA Pro Pro Lys Pro Lys Pro Ser Leu Ser Ser His Pro Thr Ser Pro Thr 2140 2145 2150	6602
GCT GGC CAG GAG CCG GGA CCC CAC CCA CAG GCC TCA GCC GTG GGC Ala Gly Gln Glu Pro Gly Pro His Pro Gln Ala Gly Ser Ala Val Gly 2155 2160 2165	6650
TTT CCG AAC ACA ACG CCC TGC TGC AGA GAG ACC CCC TCA GCC AGC CCC Phe Pro Asn Thr Thr Pro Cys Cys Arg Glu Thr Pro Ser Ala Ser Pro 2170 2175 2180 2185	6698
TGG CCC CTG GCT CTC GAA TTG GCT CTG ACC CTT ACC TGG GGC AGC GTC Trp Pro Leu Ala Leu Glu Leu Ala Leu Thr Leu Thr Trp Gly Ser Val 2190 2195 2200	6746
TGG ACA GTG AGG CCT CTG TCC ACG CCC TGC CTG AGG ACA CGC TCA CTT Trp Thr Val Arg Pro Leu Ser Thr Pro Cys Leu Arg Thr Arg Ser Leu 2205 2210 2215	6794
TCG AGG AGG CTG TGG CCA CCA ACT CGG GCC GCT CCT CCA GGA CTT CCT Ser Arg Arg Leu Trp Pro Pro Thr Arg Ala Ala Pro Pro Gly Leu Pro 2220 2225 2230	6842
ACG TGT CCT CCC TGACCTCCCCA GTCTCACCCCT CTCCGCCGCC TGCCCAACGG Thr Cys Pro Pro 2235	6894
TTACCACTGC ACCCTGGGAC TCAGCTCGGG TGGCCGAGCA CGGCACAGCT ACCACCACCC	6954
TGACCAAGAC CACTGGTGCT AGCTGCACCG TGACCGCTCA GACGCCCTGA TGCAAGCAGGC	7014
GTGTGTTCCA GTGGATGAGT TTTATCATCC ACACGGGGCA GTCGCCCTC GGGGGAGGCC	7074
TTGCCCACCT TGGTGAGGCT CCTGTGGCCC CTCCCTCCCC CTCCCTCCCT CTTTTACTCT	7134
AGACGACCAA TAAAGCCCTG TTGCTTGAGT GTACGTACCG C	7175

## (2) INFORMATION FOR SEQ ID NO:9:

## (i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 1546 base pairs
- (B) TYPE: nucleic acid

-129-

(C) STRANDEDNESS: double  
 (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (genomic)

(ix) FEATURE:

(A) NAME/KEY: CDS  
 (B) LOCATION: 1..1437

(ix) FEATURE:

(A) NAME/KEY: 3'UTR  
 (B) LOCATION: 1435..1546

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:9:

ATG GTC CAG AAG ACC AGC ATG TCC CGG CCT TAC CCA CCC TCC CAG Met Val Gln Lys Thr Ser Met Ser Arg Gly Pro Tyr Pro Pro Ser Gln 1 5 10 15	48
GAG ATC CCC ATG GAG GTC TTC GAC CCC AGC CCG CAG GGC AAA TAC AGC Glu Ile Pro Met Glu Val Phe Asp Pro Ser Pro Gln Gly Lys Tyr Ser 20 25 30	96
AAG AGG AAA GGG CGA TTC AAA CGG TCA GAT GGG AGC ACG TCC TCG GAT Lys Arg Lys Gly Arg Phe Lys Arg Ser Asp Gly Ser Thr Ser Ser Asp 35 40 45	144
ACC ACA TCC AAC AGC TTT GTC CGC CAG GGC TCA GCG GAG TCC TAC ACC Thr Thr Ser Asn Ser Phe Val Arg Gln Gly Ser Ala Glu Ser Tyr Thr 50 55 60	192
AGC CGT CCA TCA GAC TCT GAT GTA TCT CTG GAG GAC CGG GAA GCC Ser Arg Pro Ser Asp Ser Asp Val Ser Leu Glu Glu Asp Arg Glu Ala 65 70 75 80	240
TTA AGG AAG GAA GCA GAG CGC CAG GCA TTA GCG CAG CTC GAG AAG GCC Leu Arg Lys Glu Ala Glu Arg Gln Ala Leu Ala Gln Leu Glu Lys Ala 85 90 95	288
AAG ACC AAG CCA GTG GCA TTT GCT GTG CGG ACA AAT GTT GGC TAC AAT Lys Thr Lys Pro Val Ala Phe Ala Val Arg Thr Asn Val Gly Tyr Asn 100 105 110	336
CCG TCT CCA GGG GAT GAG GTG CCT GTG CAG GGA GTG GCC ATC ACC TTC Pro Ser Pro Gly Asp Glu Val Pro Val Gln Gly Val Ala Ile Thr Phe 115 120 125	384
GAG CCC AAA GAC TTC CTG CAC ATC AAG GAG AAA TAC AAT AAT GAC TGG Glu Pro Lys Asp Phe Leu His Ile Lys Glu Lys Tyr Asn Asn Asp Trp 130 135 140	432
TGG ATC GGG CGG CTG GTG AAG GAG GGC TGT GAG GTT GGC TTC ATT CCC Trp Ile Gly Arg Leu Val Lys Glu Gly Cys Glu Val Gly Phe Ile Pro 145 150 155 160	480
AGC CCC GTC AAA CTG GAC AGC CTT CGC CTG CTG CAG GAA CAG AAG CTG Ser Pro Val Lys Leu Asp Ser Leu Arg Leu Leu Gln Glu Gln Lys Leu 165 170 175	528
CGC CAG AAC CGC CTC GGC TCC AGC AAA TCA GGC GAT AAC TCC AGT TCC Arg Gln Asn Arg Leu Gly Ser Ser Lys Ser Gly Asp Asn Ser Ser Ser 180 185 190	576
AGT CTG GGA GAT GTG GTG ACT GGC ACC CGC CGC CCC ACA CCC CCT GCC Ser Leu Gly Asp Val Val Thr Gly Thr Arg Arg Pro Thr Pro Pro Ala 195 200 205	624

-130-

AGT GCC AAA CAG AAG CAG AAG TCG ACA GAG CAT GTG CCC CCC TAT GAC Ser Ala Lys Gln Lys Gln Ser Thr Glu His Val Pro Pro Tyr Asp 210 215 220	672
GTG GTG CCT TCC ATG AGG CCC ATC ATC CTG GTG GGA CCG TCG CTC AAG Val Val Pro Ser Met Arg Pro Ile Ile Leu Val Gly Pro Ser Leu Lys 225 230 235 240	720
GTC TAC GAG GTT ACA GAC ATG ATG CAG AAA GCT TTA TTT GAC TTC TTG Gly Tyr Glu Val Thr Asp Met Met Gln Lys Ala Leu Phe Asp Phe Leu 245 250 255	768
AAG CAT CGG TTT GAT GGC AGG ATC TCC ATC ACT CGT GTG ACG GCA GAT Lys His Arg Phe Asp Gly Arg Ile Ser Ile Thr Arg Val Thr Ala Asp 260 265 270	816
ATT TCC CTG GCT AAG CGC TCA GTT CTC AAC AAC CCC AGC AAA CAC ATC Ile Ser Leu Ala Lys Arg Ser Val Leu Asn Asn Pro Ser Lys His Ile 275 280 285	864
ATC ATT GAG CGC TCC AAC ACA CGC TCC AGC CTG GCT GAG GTG CAG AGT Ile Ile Glu Arg Ser Asn Thr Arg Ser Ser Leu Ala Glu Val Gln Ser 290 295 300	912
GAA ATC GAG CGA ATC TTC GAG CTG GCC CGG ACC CTT CAG TTG GTC GCT Glu Ile Glu Arg Ile Phe Glu Leu Ala Arg Thr Leu Gln Leu Val Ala 305 310 315 320	960
CTG GAT GCT GAC ACC ATC AAT CAC CCA GCC CAG CTG TCC AAG ACC TCG Leu Asp Ala Asp Thr Ile Asn His Pro Ala Gln Leu Ser Lys Thr Ser 325 330 335	1008
CTG GCC CCC ATC ATT GTT TAC ATC AAG ATC ACC TCT CCC AAG GTA CTT Leu Ala Pro Ile Ile Val Tyr Ile Lys Ile Thr Ser Pro Lys Val Leu 340 345 350	1056
CAA AGG CTC ATC AAG TCC CGA GGA AAG TCT CAG TCC AAA CAC CTC AAT Gln Arg Leu Ile Lys Ser Arg Gly Lys Ser Gln Ser Lys His Leu Asn 355 360 365	1104
GTC CAA ATA GCG GCC TCG GAA AAG CTG GCA CAG TGC CCC CCT GAA ATG Val Gln Ile Ala Ala Ser Glu Lys Leu Ala Gln Cys Pro Pro Glu Met 370 375 380	1152
TTT GAC ATC ATC CTG GAT GAG AAC CAA TTG GAG GAT GCC TGC GAG CAT Phe Asp Ile Ile Leu Asp Glu Asn Gln Leu Glu Asp Ala Cys Glu His 385 390 395 400	1200
CTG GCG GAG TAC TTG GAA GCC TAT TGG AAG GCC ACA CAC CCG CCC AGC Leu Ala Glu Tyr Leu Glu Ala Tyr Trp Lys Ala Thr His Pro Pro Ser 405 410 415	1248
AGC ACG CCA CCC AAT CCG CTG CTG AAC CGC ACC ATG GCT ACC GCA GCC Ser Thr Pro Pro Asn Pro Leu Leu Asn Arg Thr Met Ala Thr Ala Ala 420 425 430	1296
CTG GCT GCC AGC CCT GCC CCT GTC TCC AAC CTC CAG GTA CAG GTG CTC Leu Ala Ala Ser Pro Ala Pro Val Ser Asn Leu Gln Val Gln Val Leu 435 440 445	1344
ACC TCG CTC AGG AGA AAC CTC GGC TTC TGG GGC GGG CTG GAG TCC TCA Thr Ser Leu Arg Arg Asn Leu Gly Phe Trp Gly Gly Leu Glu Ser Ser 450 455 460	1392
CAG CGG GGC AGT GTG GTG CCC CAG GAG CAG GAA CAT GCC ATG TAGTGGCGC Gln Arg Gly Ser Val Val Pro Gln Glu Gln Glu His Ala Met 465 470 475	1444

-131-

CCTGCCCGTC TTCCCTCCGT CTCTGGGTC GGAAGTGGAG TGCAGGGAAC ATGGAGGAGG	1504
AAGGGAAAGAG CTTTATTTG TAAAAAAAATA AGATGAGCGG CA	1546

## (2) INFORMATION FOR SEQ ID NO:10:

## (i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 1851 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: double
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (genomic)

## (ix) FEATURE:

- (A) NAME/KEY: CDS
- (B) LOCATION: 1..1797
- (D) OTHER INFORMATION: /standard\_name= "Beta-3"

## (ix) FEATURE:

- (A) NAME/KEY: 3'UTR
- (B) LOCATION: 1795..1851

## (xi) SEQUENCE DESCRIPTION: SEQ ID NO:10:

ATG GTC CAG AAG ACC AGC ATG TCC CGG GGC CCT TAC CCA CCC TCC CAG Met Val Gln Lys Thr Ser Met Ser Arg Gly Pro Tyr Pro Pro Ser Gln	48
1 5 10 15	
GAG ATC CCC ATG GGA GTC TTC GAC CCC AGC CCG CAG GGC AAA TAC AGC Glu Ile Pro Met Gly Val Phe Asp Pro Ser Pro Gln Gly Lys Tyr Ser	96
20 25 30	
AAG AGG AAA GGG CGA TTC AAA CGG TCA GAT GGG AGC ACG TCC TCG GAT Lys Arg Lys Gly Arg Phe Lys Arg Ser Asp Gly Ser Thr Ser Ser Asp	144
35 40 45	
ACC ACA TCC AAC AGC TTT GTC CGC CAG GGC TCA GCG GAG TCC TAC ACC Thr Thr Ser Asn Ser Phe Val Arg Gln Gly Ser Ala Glu Ser Tyr Thr	192
50 55 60	
AGC CGT CCA TCA GAC TCT GAT GTA TCT CTG GAG GAG GAC CGG GAA GCC Ser Arg Pro Ser Asp Ser Asp Val Ser Leu Glu Glu Asp Arg Glu Ala	240
65 70 75 80	
TTA AGG AAG GAA GCA GAG CGC CAG GCA TTA GCG CAG CTC GAG AAG GCC Leu Arg Lys Glu Ala Glu Arg Gln Ala Leu Ala Gln Leu Glu Lys Ala	288
85 90 95	
AAG ACC AAG CCA GTG GCA TTT GCT GTG CGG ACA AAT GTT GGC TAC AAT Lys Thr Lys Pro Val Ala Phe Ala Val Arg Thr Asn Val Gly Tyr Asn	336
100 105 110	
CCG TCT CCA GGG GAT GAG GTG CCT GTG CAG GGA GTG GCC ATC ACC TTC Pro Ser Pro Gly Asp Glu Val Pro Val Gln Gly Val Ala Ile Thr Phe	384
115 120 125	
GAG CCC AAA GAC TTC CTG CAC ATC AAG GAG AAA TAC AAT AAT GAC TGG Glu Pro Lys Asp Ph Leu His Ile Lys Glu Lys Tyr Asn Asn Asp Trp	432
130 135 140	
TGG ATC GGG CGG CTG GTG AAG GAG GGC TGT GAG GTT GGC TTC ATT CCC Trp Ile Gly Arg Leu Val Lys Glu Gly Cys Glu Val Gly Phe Ile Pro	480
145 150 155 160	

-132-

AGC CCC GTC AAA CTG GAC AGC CTT CGC CTG CTG CAG GAA CAG AAG CTG Ser Pro Val Lys Leu Asp Ser Leu Arg Leu Leu Gln Glu Gln Lys Leu 165 170 175	528
CGC CAG AAC CGC CTC GGC TCC AGC AAA TCA GGC GAT AAC TCC AGT TCC Arg Gln Asn Arg Leu Gly Ser Ser Lys Ser Gly Asp Asn Ser Ser Ser 180 185 190	576
AGT CTG GGA GAT GTG GTG ACT GGC ACC CGC CGC CCC ACA CCC CCT GCC Ser Leu Gly Asp Val Val Thr Gly Thr Arg Arg Pro Thr Pro Pro Ala 195 200 205	624
AGT GCC AAA CAG AAG CAG AAG TCG ACA GAG CAT GTG CCC CCC TAT GAC Ser Ala Lys Gln Lys Gln Lys Ser Thr Glu His Val Pro Pro Tyr Asp 210 215 220	672
GTG GTG CCT TCC ATG AGG CCC ATC ATC CTG GTG GGA CCG TCG CTC AAG Val Val Pro Ser Met Arg Pro Ile Ile Leu Val Gly Pro Ser Leu Lys 225 230 235 240	720
GGC TAC GAG GTT ACA GAC ATG ATG CAG AAA GCT TTA TTT GAC TTC TTG Gly Tyr Glu Val Thr Asp Met Met Gln Lys Ala Leu Phe Asp Phe Leu 245 250 255	768
AAG CAT CGG TTT GAT GGC AGG ATC TCC ATC ACT CGT GTG ACG GCA GAT Lys His Arg Phe Asp Gly Arg Ile Ser Ile Thr Arg Val Thr Ala Asp 260 265 270	816
ATT TCC CTG GCT AAG CGC TCA GTT CTC AAC AAC CCC AGC AAA CAC ATC Ile Ser Leu Ala Lys Arg Ser Val Leu Asn Asn Pro Ser Lys His Ile 275 280 285	864
ATC ATT GAG CGC TCC AAC ACA CGC TCC AGC CTG GCT GAG GTG CAG AGT Ile Ile Glu Arg Ser Asn Thr Arg Ser Ser Leu Ala Glu Val Gln Ser 290 295 300	912
GAA ATC GAG CGA ATC TTC GAG CTG GCC CGG ACC CTT CAG TTG GTC GCT Glu Ile Glu Arg Ile Phe Glu Leu Ala Arg Thr Leu Gln Leu Val Ala 305 310 315 320	960
CTG GAT GCT GAC ACC ATC AAT CAC CCA GCC CAG CTG TCC AAG ACC TCG Leu Asp Ala Asp Thr Ile Asn His Pro Ala Gln Leu Ser Lys Thr Ser 325 330 335	1008
CTG GCC CCC ATC ATT GTT TAC ATC AAG ATC ACC TCT CCC AAG GTA CTT Leu Ala Pro Ile Ile Val Tyr Ile Lys Ile Thr Ser Pro Lys Val Leu 340 345 350	1056
CAA AGG CTC ATC AAG TCC CGA GGA AAG TCT CAG TCC AAA CAC CTC AAT Gln Arg Leu Ile Lys Ser Arg Gly Lys Ser Gln Ser Lys His Leu Asn 355 360 365	1104
GTC CAA ATA GCG GCC TCG GAA AAG CTG GCA CAG TGC CCC CCT GAA ATG Val Gln Ile Ala Ala Ser Glu Lys Leu Ala Gln Cys Pro Pro Glu Met 370 375 380	1152
TTT GAC ATC ATC CTG GAT GAG AAC CAA TTG GAG GAT GCC TGC GAG CAT Phe Asp Ile Ile Leu Asp Glu Asn Gln Leu Glu Asp Ala Cys Glu His 385 390 395 400	1200
CTG GCG GAG TAC TTG GAA TAT TGG AAG GCC ACA CAC CCG CCC AGC Leu Ala Glu Tyr Leu Glu Ala Tyr Trp Lys Ala Thr His Pro Pro Ser 405 410 415	1248
AGC ACG CCA CCC AAT CCG CTG AAC CGC ACC ATG GCT ACC GCA GCC Ser Thr Pro Pro Asn Pro Leu Leu Asn Arg Thr Met Ala Thr Ala Ala 420 425 430	1296

-133-

CTG GCT GCC AGC CCT GCC CCT GTC TCC AAC CTC CAG GGA CCC TAC CTT Leu Ala Ala Ser Pro Ala Pro Val Ser Asn Leu In Gly Pro Tyr Leu 435 440 445	1344
GCT TCC GGG GAC CAG CCA CTG GAA CGG GCC ACC GGG GAG CAC GCC AGC Ala Ser Gly Asp Gln Pro Leu Glu Arg Ala Thr Gly Glu His Ala Ser 450 455 460	1392
ATG CAC GAG TAC CCA GGG GAG CTG GGC CAG CCC CCA GCC CTT TAC CCC Met His Glu Tyr Pro Gly Glu Leu Gly Gln Pro Pro Gly Leu Tyr Pro 465 470 475 480	1440
AGC AGC CAC CCA CCA GGC CGG GCA GGC ACG CTA CGG GCA CTG TCC CGC Ser Ser His Pro Pro Gly Arg Ala Gly Thr Leu Arg Ala Leu Ser Arg 485 490 495	1488
CAA GAC ACT TTT GAT GCC GAC ACC CCC GGC AGC CGA AAC TCT GCC TAC Gln Asp Thr Phe Asp Ala Asp Thr Pro Gly Ser Arg Asn Ser Ala Tyr 500 505 510	1536
ACG GAG CTG GGA GAC TCA TGT GTG GAC ATG GAG ACT GAC CCC TCA GAG Thr Glu Leu Gly Asp Ser Cys Val Asp Met Glu Thr Asp Pro Ser Glu 515 520 525	1584
GGG CCA GGG CTT GGA GAC CCT GCA GGG GGC GGC ACG CCC CCA GCC CGA Gly Pro Gly Leu Gly Asp Pro Ala Gly Gly Thr Pro Pro Ala Arg 530 535 540	1632
CAG GGA TCC TGG GAG GAC GAG GAA GAA GAC TAT GAG GAA GAG CTG ACC Gln Gly Ser Trp Glu Asp Glu Glu Asp Tyr Glu Glu Glu Leu Thr 545 550 555 560	1680
GAC AAC CGG AAC CGG GGC CGG AAT AAG GCC CGC TAC TGC GCT GAG GGT Asp Asn Arg Asn Arg Gly Arg Asn Lys Ala Arg Tyr Cys Ala Glu Gly 565 570 575	1728
GGG GGT CCA GTT TTG GGG CGC AAC AAG AAT GAG CTG GAG GGC TGG GGA Gly Gly Pro Val Leu Gly Arg Asn Lys Asn Glu Leu Glu Gly Trp Gly 580 585 590	1776
CGA GGC GTC TAC ATT CGC TGAGAGGCAG GGGCCACACG GCGGGAGGAA Arg Gly Val Tyr Ile Arg 595	1824
GGGCTCTGAG CCCAGGGGAG GGGAGGG	1851

## (2) INFORMATION FOR SEQ ID NO:11:

## (i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 3600 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: double
- (D) TOPOLOGY: linear

## (ii) MOLECULE TYPE: DNA (genomic)

## (ix) FEATURE:

- (A) NAME/KEY: CDS
- (B) LOCATION: 35..3310
- (D) OTHER INFORMATION: /standard\_name= "Alpha-2"

## (ix) FEATURE:

- (A) NAME/KEY: 5'UTR
- (B) LOCATION: 1..34

## (ix) FEATURE:

- (A) NAME/KEY: 3'UTR

-134-

(B) LOCATI N: 3308..3600

(xi) SEQUENCE DESCRIPTION: SEQ ID N :11:

GCAGGGGGAGG GGGCATTGAT CTTCGATCGC GAAG ATG GCT GCT GGC TGC CTG Met Ala Ala Gly Cys Leu	52
1 5	
CTG GCC TTG ACT CTG ACA CTT TTC CAA TCT TTG CTC ATC GGC CCC TCG Leu Ala Leu Thr Leu Thr Leu Phe Gln Ser Leu Leu Ile Gly Pro Ser	100
10 15 20	
TCG GAG GAG CCG TTC CCT TCG GCC GTC ACT ATC AAA TCA TGG GTG GAT Ser Glu Glu Pro Phe Pro Ser Ala Val Thr Ile Lys Ser Trp Val Asp	148
25 30 35	
AAG ATG CAA GAA GAC CTT GTC ACA CTG GCA AAA ACA GCA AGT GGA GTC Lys Met Gln Glu Asp Leu Val Thr Leu Ala Lys Thr Ala Ser Gly Val	196
40 45 50	
AAT CAG CTT GTT GAT ATT TAT GAG AAA TAT CAA GAT TTG TAT ACT GTG Asn Gln Leu Val Asp Ile Tyr Glu Lys Tyr Gln Asp Leu Tyr Thr Val	244
55 60 65 70	
GAA CCA AAT AAT GCA CGC CAG CTG GTC GAA ATT GCA GCC AGG GAT ATT Glu Pro Asn Asn Ala Arg Gln Leu Val Glu Ile Ala Ala Arg Asp Ile	292
75 80 85	
GAG AAA CTT CTG AGC AAC AGA TCT AAA GCC CTG GTG AGC CTG GCA TTG Glu Lys Leu Leu Ser Asn Arg Ser Lys Ala Leu Val Ser Leu Ala Leu	340
90 95 100	
GAA GCG GAG AAA GTT CAA GCA GCT CAC CAG TGG AGA GAA GAT TTT GCA Glu Ala Glu Lys Val Gln Ala Ala His Gln Trp Arg Glu Asp Phe Ala	388
105 110 115	
AGC AAT GAA GTT GTC TAC TAC AAT GCA AAG GAT GAT CTC GAT CCT GAG Ser Asn Glu Val Val Tyr Tyr Asn Ala Lys Asp Asp Leu Asp Pro Glu	436
120 125 130	
AAA AAT GAC AGT GAG CCA GGC AGC CAG AGG ATA AAA CCT GTT TTC ATT Lys Asn Asp Ser Glu Pro Gly Ser Gln Arg Ile Lys Pro Val Phe Ile	484
135 140 145 150	
GAA GAT GCT AAT TTT GGA CGA CAA ATA TCT TAT CAG CAC GCA GCA GTC Glu Asp Ala Asn Phe Gly Arg Gln Ile Ser Tyr Gln His Ala Ala Val	532
155 160 165	
CAT ATT CCT ACT GAC ATC TAT GAG GGC TCA ACA ATT GTG TTA AAT GAA His Ile Pro Thr Asp Ile Tyr Glu Gly Ser Thr Ile Val Leu Asn Glu	580
170 175 180	
CTC AAC TGG ACA AGT GCC TTA GAT GAA GTT TTC AAA AAG AAT CGC GAG Leu Asn Trp Thr Ser Ala Leu Asp Glu Val Phe Lys Lys Asn Arg Glu	628
185 190 195	
GAA GAC CCT TCA TTA TTG TGG CAG GTT TTT GGC AGT GCC ACT GGC CTA Glu Asp Pro Ser Leu Leu Trp Gln Val Phe Gly Ser Ala Thr Gly Leu	676
200 205 210	
GCT CGA TAT TAT CCA GCT TCA CCA TGG GTT GAT AAT AGT AGA ACT CCA Ala Arg Tyr Tyr Pro Ala Ser Pro Trp Val Asp Asn Ser Arg Thr Pro	724
215 220 225 230	
AAT AAG ATT GAC CTT TAT GAT GTA CGC AGA AGA CCA TGG TAC ATC CAA Asn Lys Ile Asp Leu Tyr Asp Val Arg Arg Arg Pro Trp Tyr Ile Gin	772
235 240 245	

-135-

GGA GCT GCA TCT CCT AAA GAC ATG CTT ATT CTG GTG GAT GTG AGT GGA Gly Ala Ala Ser Pro Lys Asp Met Leu Ile Leu Val Asp Val Ser Gly 250 255 260	820
AGT GTT AGT GGA TTG ACA CTT AAA CTG ATC CGA ACA TCT GTC TCC GAA Ser Val Ser Gly Leu Thr Leu Lys Leu Ile Arg Thr Ser Val Ser Glu 265 270 275	868
ATG TTA GAA ACC CTC TCA GAT GAT GAT TTC GTG AAT GTA GCT TCA TTT Met Leu Glu Thr Leu Ser Asp Asp Asp Phe Val Asn Val Ala Ser Phe 280 285 290	916
AAC AGC AAT GCT CAG GAT GTA AGC TGT TTT CAG CAC CTT GTC CAA GCA Asn Ser Asn Ala Gln Asp Val Ser Cys Phe Gln His Leu Val Gln Ala 295 300 305 310	964
AAT GTA AGA AAT AAA AAA GTG TTG AAA GAC GCG GTG AAT AAT ATC ACA Asn Val Arg Asn Lys Lys Val Leu Lys Asp Ala Val Asn Asn Ile Thr 315 320 325	1012
GCC AAA GGA ATT ACA GAT TAT AAG AAG GGC TTT AGT TTT GCT TTT GAA Ala Lys Gly Ile Thr Asp Tyr Lys Lys Gly Phe Ser Phe Ala Phe Glu 330 335 340	1060
CAG CTG CTT AAT TAT AAT GTT TCC AGA GCA AAC TGC AAT AAG ATT ATT Gln Leu Leu Asn Tyr Asn Val Ser Arg Ala Asn Cys Asn Lys Ile Ile 345 350 355	1108
ATG CTA TTC ACG GAT GGA GGA GAA GAG AGA GCC CAG GAG ATA TTT AAC Met Leu Phe Thr Asp Gly Gly Glu Glu Arg Ala Gln Glu Ile Phe Asn 360 365 370	1156
AAA TAC AAT AAA GAT AAA AAA GTA CGT GTA TTC AGG TTT TCA GTT GGT Lys Tyr Asn Lys Asp Lys Lys Val Arg Val Phe Arg Phe Ser Val Gly 375 380 385 390	1204
CAA CAC AAT TAT GAG AGA GGA CCT ATT CAG TGG ATG GCC TGT GAA AAC Gln His Asn Tyr Glu Arg Gly Pro Ile Gln Trp Met Ala Cys Glu Asn 395 400 405	1252
AAA GCT TAT TAT GAA ATT CCT TCC ATT CGT GCA ATA AGA ATC AAT Lys Gly Tyr Tyr Glu Ile Pro Ser Ile Gly Ala Ile Arg Ile Asn 410 415 420	1300
ACT CAG GAA TAT TTG GAT GTT TTG GGA AGA CCA ATG GTT TTA GCA GGA Thr Gln Glu Tyr Leu Asp Val Leu Gly Arg Pro Met Val Leu Ala Gly 425 430 435	1348
GAC AAA GCT AAG CAA GTC CAA TGG ACA AAT GTG TAC CTG GAT GCA TTG Asp Lys Ala Lys Gln Val Gln Trp Thr Asn Val Tyr Leu Asp Ala Leu 440 445 450	1396
GAA CTG GGA CTT GTC ATT ACT GGA ACT CTT CCG GTC TTC AAC ATA ACC Glu Leu Gly Leu Val Ile Thr Gly Thr Leu Pro Val Phe Asn Ile Thr 455 460 465 470	1444
GGC CAA TTT GAA AAT AAG ACA AAC TTA AAG AAC CAG CTG ATT CTT GGT Gly Gln Phe Glu Asn Lys Thr Asn Leu Lys Asn Gln Leu Ile Leu Gly 475 480 485	1492
GTG ATG GGA GTA GAT GTG TCT TTG GAA GAT ATT AAA AGA CTG ACA CCA Val Met Gly Val Asp Val Ser Leu Glu Asp Ile Lys Arg Leu Thr Pro 490 495 500	1540
CGT TTT ACA CTG TGC CCC AAT GGG TAT TAC TTT GCA ATC GAT CCT AAT Arg Phe Thr Leu Cys Pro Asn Gly Tyr Tyr Phe Ala Ile Asp Pro Asn 505 510 515	1588

-136-

GGT TAT GTT TTA TTA CAT CCA AAT CTT CAG CCA AAG AAC CCC AAA TCT Gly Tyr Val Leu Leu His Pro Asn Leu Gln Pro Lys Asn Pro Lys Ser 520 525 530	1636
CAG GAG CCA GTA ACA TTG GAT TTC CTT GAT GCA GAG TTA GAG AAT GAT Gln Glu Pro Val Thr Leu Asp Phe Leu Asp Ala Glu Leu Glu Asn Asp 535 540 545 550	1684
ATT AAA GTG GAG ATT CGA AAT AAG ATG ATT GAT GGG GAA AGT GGA GAA Ile Lys Val Glu Ile Arg Asn Lys Met Ile Asp Gly Glu Ser Gly Glu 555 560 565	1732
AAA ACA TTC AGA ACT CTG GTT AAA TCT CAA GAT GAG AGA TAT ATT GAC Lys Thr Phe Arg Thr Leu Val Lys Ser Gln Asp Glu Arg Tyr Ile Asp 570 575 580	1780
AAA GGA AAC AGG ACA TAC ACA TGG ACA CCT GTC AAT GGC ACA GAT TAC Lys Gly Asn Arg Thr Tyr Thr Trp Thr Pro Val Asn Gly Thr Asp Tyr 585 590 595	1828
AGT TTG GCC TTG GTA TTA CCA ACC TAC AGT TTT TAC TAT ATA AAA GCC Ser Leu Ala Leu Val Leu Pro Thr Tyr Ser Phe Tyr Tyr Ile Lys Ala 600 605 610	1876
AAA CTA GAA GAG ACA ATA ACT CAG GCC AGA TCA AAA AAG GGC AAA ATG Lys Leu Glu Glu Thr Ile Thr Gln Ala Arg Ser Lys Lys Gly Lys Met 615 620 625 630	1924
AAG GAT TCG GAA ACC CTG AAG CCA GAT AAT TTT GAA GAA TCT GGC TAT Lys Asp Ser Glu Thr Leu Lys Pro Asp Asn Phe Glu Glu Ser Gly Tyr 635 640 645	1972
ACA TTC ATA GCA CCA AGA GAT TAC TGC AAT GAC CTG AAA ATA TCG GAT Thr Phe Ile Ala Pro Arg Asp Tyr Cys Asn Asp Leu Lys Ile Ser Asp 650 655 660	2020
AAT AAC ACT GAA TTT CTT TTA AAT TTC AAC GAG TTT ATT GAT AGA AAA Asn Asn Thr Glu Phe Leu Leu Asn Phe Asn Glu Phe Ile Asp Arg Lys 665 670 675	2068
ACT CCA AAC AAC CCA TCA TGT AAC GCG GAT TTG ATT AAT AGA GTC TTG Thr Pro Asn Asn Pro Ser Cys Asn Ala Asp Leu Ile Asn Arg Val Leu 680 685 690	2116
CTT GAT GCA GGC TTT ACA AAT GAA CTT GTC CAA AAT TAC TGG AGT AAG Leu Asp Ala Gly Phe Thr Asn Glu Leu Val Gln Asn Tyr Trp Ser Lys 695 700 705 710	2164
CAG AAA AAT ATC AAG GGA GTG AAA GCA CGA TTT GTT GTG ACT GAT GGT Gln Lys Asn Ile Lys Gly Val Lys Ala Arg Phe Val Val Thr Asp Gly 715 720 725	2212
GGG ATT ACC AGA GTT TAT CCC AAA GAG GCT GGA GAA AAT TGG CAA GAA Gly Ile Thr Arg Val Tyr Pro Lys Glu Ala Gly Glu Asn Trp Gln Glu 730 735 740	2260
AAC CCA GAG ACA TAT GAG GAC AGC TTC TAT AAA AGG AGC CTA GAT AAT Asn Pro Glu Thr Tyr Glu Asp Ser Phe Tyr Lys Arg Ser Leu Asp Asn 745 750 755	2308
GAT AAC TAT GTT TTC ACT GCT CCC TAC TTT AAC AAA AGT GGA CCT GGT Asp Asn Tyr Val Phe Thr Ala Pro Tyr Phe Asn Lys Ser Gly Pro Gly 760 765 770	2356
GCC TAT GAA TCG GGC ATT ATG GTA AGC AAA GCT GTA GAA ATA TAT ATT Ala Tyr Glu Ser Gly Ile Met Val Ser Lys Ala Val Glu Ile Tyr Ile 775 780 785 790	2404

-137-

CAA GGG AAA CTT CTT AAA CCT GCA GTT GTT GGA ATT AAA ATT GAT GTA Gln Gly Lys Leu Leu Lys Pro Ala Val Val Gly Ile Lys Ile Asp Val 795 800 805	2452
AAT TCC TGG ATA GAG AAT TTC ACC AAA ACC TCA ATC AGA GAT CCG TGT Asn Ser Trp Ile Glu Asn Phe Thr Lys Thr Ser Ile Arg Asp Pro Cys 810 815 820	2500
GCT GGT CCA GTT TGT GAC TGC AAA AGA AAC AGT GAC GTA ATG GAT TGT Ala Gly Pro Val Cys Asp Cys Lys Arg Asn Ser Asp Val Met Asp Cys 825 830 835	2548
GTG ATT CTG GAT GAT GGT GGG TTT CTT CTG ATG GCA AAT CAT GAT GAT Val Ile Leu Asp Asp Gly Gly Phe Leu Leu Met Ala Asn His Asp Asp 840 845 850	2596
TAT ACT AAT CAG ATT GGA AGA TTT TTT GGA GAG ATT GAT CCC AGC TTG Tyr Thr Asn Gln Ile Gly Arg Phe Phe Gly Glu Ile Asp Pro Ser Leu 855 860 865 870	2644
ATG AGA CAC CTG GTT AAT ATA TCA GTT TAT GCT TTT AAC AAA TCT TAT Met Arg His Leu Val Asn Ile Ser Val Tyr Ala Phe Asn Lys Ser Tyr 875 880 885	2692
GAT TAT CAG TCA GTA TGT GAG CCC GGT GCT GCA CCA AAA CAA GGA GCA Asp Tyr Gln Ser Val Cys Glu Pro Gly Ala Ala Pro Lys Gln Gly Ala 890 895 900	2740
GGA CAT CGC TCA GCA TAT GTG CCA TCA GTA GCA GAC ATA TTA CAA ATT Gly His Arg Ser Ala Tyr Val Pro Ser Val Ala Asp Ile Leu Gln Ile 905 910 915	2788
GGC TGG TGG GCC ACT GCT GCT GCC TGG TCT ATT CTA CAG CAG TTT CTC Gly Trp Trp Ala Thr Ala Ala Trp Ser Ile Leu Gln Gln Phe Leu 920 925 930	2836
TTG AGT TTG ACC TTT CCA CGA CTC CTT GAG GCA GTT GAG ATG GAG GAT Leu Ser Leu Thr Phe Pro Arg Leu Leu Glu Ala Val Glu Met Glu Asp 935 940 945 950	2884
GAT GAC TTC ACG GCC TCC CTG TCC AAG CAG AGC TGC ATT ACT GAA CAA Asp Asp Phe Thr Ala Ser Leu Ser Lys Gln Ser Cys Ile Thr Glu Gln 955 960 965	2932
ACC CAG TAT TTC TTC GAT AAC GAC AGT AAA TCA TTC AGT GGT GTA TTA Thr Gln Tyr Phe Phe Asp Asn Asp Ser Lys Ser Phe Ser Gly Val Leu 970 975 980	2980
GAC TGT GGA AAC TGT TCC AGA ATC TTT CAT GGA GAA AAG CCT ATG AAC Asp Cys Gly Asn Cys Ser Arg Ile Phe His Gly Glu Lys Leu Met Asn 985 990 995	3028
ACC AAC TTA ATA TTC ATA ATG GTT GAG AGC AAA GGG ACA TGT CCA TGT Thr Asn Leu Ile Phe Ile Met Val Glu Ser Lys Gly Thr Cys Pro Cys 1000 1005 1010	3076
GAC ACA CGA CTG CTC ATA CAA GCG GAG CAG ACT TCT GAC GGT CCA AAT Asp Thr Arg Leu Leu Ile Gln Ala Glu Gln Thr Ser Asp Gly Pro Asn 1015 1020 1025 1030	3124
CCT TGT GAC ATG GTT AAG CAA CCT AGA TAC CGA AAA GGG CCT GAT GTC Pro Cys Asp Met Val Lys Gln Pro Arg Tyr Arg Lys Gly Pro Asp Val 1035 1040 1045	3172
TGC TTT GAT AAC AAT GTC TTG GAG GAT TAT ACT GAC TGT GGT GGT GTT Cys Phe Asp Asn Asn Val Leu Glu Asp Tyr Thr Asp Cys Gly Gly Val 1050 1055 1060	3220

-138-

TCT GGA TTA AAT CCC TCC CTG TGG TAT ATC ATT GGA ATC CAG TTT CTA	3268
Ser Gly Leu Asn Pro Ser Leu Trp Tyr Ile Ile Gly Ile Gln Phe Leu	
1065 1070 1075	
CTA CTT TGG CTG GTA TCT GGC AGC ACA CAC CGG CTG TTA TGACCTTCTA	3317
Leu Leu Trp Leu Val Ser Gly Ser Thr His Arg Leu Leu	
1080 1085 1090	
AAAACCAAAT CTGCATAGTT AAACCTCCAGA CCCTGCCAAA ACATGAGCCC TGCCCTCAAT	3377
TACAGTAACG TAGGGTCAGC TATAAAATCA GACAAACATT AGCTGGGCCT GTTCCATGGC	3437
ATAAACACTAA GGCGCAGACT CCTAAGGCAC CCACTGGCTG CATGTCAAGGG TGTCAAGATCC	3497
TTAAACGTGT GTGAATGCTG CATCATCTAT GTGTAACATC AAAGCAAAAT CCTATACGTG	3557
TCCTCTATTG GAAAATTTGG GCGTTTGTG TTGCATTGTT GGT	3600

## (2) INFORMATION FOR SEQ ID NO:12:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 323 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: double
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (genomic)

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:12:

CCCCCTGCCA GTGGCCAAAC AGAACCGAGAA GTCGGGTAAT GAAATGACTA ACTTAGCCTT	60.
TGAACTAGAC CCCCTAGAGT TAGAGGAGGA AGAGGCTGAG CTTGGTGAGC AGAGTGGCTC	120
TGCCAAGACT AGTGTAGCA GTGTACCCAC CCCGCCACCC CATGGCAAAC GCATCCCCTT	180
CTTTAAGAAG ACAGAGCATG TGCCCCCTA TGACGTGGTG CCTTCCATGA GGCCCATCAT	240
CCTGGTGGGA CCGTCGCTCA AGGGCTACGA GGTTACAGAC ATGATGCAGA AAGCTTTATT	300
TGACTTCTTG AAGCATCGGT TTG	323

## (2) INFORMATION FOR SEQ ID NO:13:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 57 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: double
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (genomic)

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:13:

CCTATTGGTG TAGGTATACC AACAAATTAAT TTAAGAAAAA GGAGACCCAA TATCCAG	57
---	----

## (2) INFORMATION FOR SEQ ID NO:14:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 180 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: double
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (genomic)

-139-

## (ix) FEATURE:

- (A) NAME/KEY: CDS
- (B) LOCATION: 1..132

## (xi) SEQUENCE DESCRIPTION: SEQ ID NO:14:

TGG TCC TTT GCC TGC GCC TGT GCC GCC	TTC ATC CTC CTC TTT CTC GGC	48
Trp Ser Phe Ala Cys Ala Cys Ala	Phe Ile Leu Leu Phe Leu Gly	
1	5	10
		15

GGT CTC GCC CTC CTG CTG TTC TCC CTG CCT CGA ATG CCC CGG AAC CCA	96	
Gly Leu Ala Leu Leu Phe Ser Leu Pro Arg Met Pro Arg Asn Pro		
20	25	30

TGG GAG TCC TGC ATG GAT GCT GAG CCC GAG CAC TAACCCTCCT GCGGCCCTAG	149
Trp Glu Ser Cys Met Asp Ala Glu Pro Glu His	
35	40

CGACCCCTCAG GCTTCTTCCC AGGAAGCGGG G 180

## (2) INFORMATION FOR SEQ ID NO:15:

## (i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 22 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: Other nucleic acid;  
(A) DESCRIPTION: Oligonucleotide

## (xi) SEQUENCE DESCRIPTION: SEQ ID NO:15:

AATTCGGTAC GTACACTCGA GC 22

## (2) INFORMATION FOR SEQ ID NO:16:

## (i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 18 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: Other nucleic acid;  
(A) DESCRIPTION: Oligonucleotide

## (xi) SEQUENCE DESCRIPTION: SEQ ID NO:16:

GCTCGAGTGT ACGTACCG 18

## (2) INFORMATION FOR SEQ ID NO:17:

## (i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 20 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: Other nucleic acid;  
(A) DESCRIPTION: Oligonucleotide

## (xi) SEQUENCE DESCRIPTION: SEQ ID NO:17:

CCATGGTACC TTGGTTGACG 20

-140-

## (2) INFORMATION FOR SEQ ID NO:18:

## (i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 24 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: Other nucleic acid;  
 (A) DESCRIPTION: Oligonucleotide

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:18:

AATTCGTCAA CGAAGGTACC ATGG

24

## (2) INFORMATION FOR SEQ ID NO:19:

## (i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 249 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: double
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: cDNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:19:

CGGTACGTAC ACTCGAGCGA CTGTGTCATG ATGTTTATTA GTGATGACAG TGAGGTGAGG	60
CAGGGGCTTG TGGAGCATGC TCTGTAGGTC ACACACTAGA GCCATAAGGC AAGAGTAGGC	120
GGGGAGACAG GTCCTCTGTG CCCTGTCTCT CCCCATCTAA CCCTAACCTA ACAAGCGGTA	180
GTTATGAGTC AGGGAACAAAC GTCTGGAGCC CCGTCCTCCA AAGATGTTTG AGGGACAAGA	240
ACAGAAATG	249

## (2) INFORMATION FOR SEQ ID NO:20:

## (i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 402 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: double
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: cDNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:20:

CCACCTAGCA CGGGTTCGTT CCCCTCTCCC GGCTGGCCC GGGCTCCCCG GTGGCCGCCG	60
CCCCCTCGCC GCCCATCTCT GGACTCAGAC GTCTCCCTGG AGGAGGACCG GGAGAGTGCC	120
CGGGGTGAAG TAGAGAGCCA GGCTCAGCAG CAGCTCGAAA GGGCCAAGCA CAAACCTGTG	180
GCATTTGCCG TGAGGGACCAA TGTCAGCTAC TGTGGCGTAC TGGATGAGGA GTGCCAGTC	240
CAGGGCTCTG GAGTCAACTT TGAGGCCAAA GATTTCTGC ACATTAAGA GAAGTACAGC	300
AATGACTGGT GGATCGGGCG GCTAGTGAAA GAGGGCGGGG ACATCGCCTT CATCCCCAGC	360
CCCCAGTGCC TGGTGAGCAT CCGCTCAAAC AGGAGCAGAA GG	402

-141-

(2) INFORMATION FOR SEQ ID NO:21:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 28 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: Other nucleic acid;  
(A) DESCRIPTION: Oligonucleotide

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:21:

CTCAGTACCA TCTCTGATAC CAGCCCCA

28

-142-

**WHAT IS CLAIMED IS:**

1. An isolated DNA fragment, comprising a sequence of nucleotides that encodes an  $\alpha_1$ -subunit of a human calcium channel.
- 5 2. The DNA fragment of claim 1, wherein the  $\alpha_1$ -subunit is a human neural calcium channel  $\alpha_1$  subunit.
3. The DNA fragment of claim 1, wherein the  $\alpha_1$  subunit is an  $\alpha_{1D}$  type  $\alpha_1$ -subunit,  $\alpha_{1C}$  type  $\alpha_1$ -subunit,  $\alpha_{1B}$  type  $\alpha_1$ -subunit or an  $\alpha_{1A}$  type  $\alpha_1$ -subunit.
- 10 4. An isolated DNA fragment, comprising a sequence of nucleotides that encodes an  $\alpha_2$ -subunit of a human calcium channel.
5. The DNA fragment of claim 4, wherein the human calcium channel is a human neural calcium channel, a human skeletal muscle calcium channel or a human aortic calcium channel.
- 15 6. The DNA fragment of claim 4, wherein the  $\alpha_2$  subunit is produced by alternative processing of a primary transcript that includes DNA encoding the amino acids set forth in sequence ID No. 11 and the DNA of Sequence ID No. 13 inserted between nucleotides 1624 and 1625 of Sequence ID No. 11.
- 20 7. The DNA fragment of claim 4, wherein the  $\alpha_2$  subunit is an  $\alpha_{2a}$ ,  $\alpha_{2b}$ ,  $\alpha_{2c}$ ,  $\alpha_{2d}$ , or an  $\alpha_{2e}$  subunit.
- 25 8. An isolated DNA fragment, comprising a sequence of nucleotides that encodes a  $\beta$ -subunit of a human calcium channel.
9. The DNA fragment of claim 8, wherein the subunit is a  $\beta_1$  or  $\beta_3$  subunit.
- 30 10. The DNA fragment of claim 8, wherein the  $\beta$  subunit is produced by alternative processing of a primary transcript that includes DNA encoding the amino acids set forth in sequence ID No. 9, but including the DNA set forth in Sequence ID No. 12 inserted in place of nucleotides 35 615-781 of Sequence ID No. 9.
11. The DNA fragment of claim 8, wherein the subunit is encoded by a transcript that lacks one or more sequences

-143-

of nucleotides selected from the group consisting of nucleotides 14-34 of Sequence ID No. 12, nucleotides 13-34 of Sequence ID No. 12, nucleotides 35-55 of Sequence ID No 12, nucleotides 56-190 of Sequence ID No. 12 and 5 nucleotides 191-271 of Sequence ID No. 12.

12. An isolated DNA fragment, comprising a sequence of nucleotides that encodes a  $\gamma$ -subunit of a human calcium channel.

13. A eukaryotic cell, comprising heterologous DNA 10 that encodes at least one subunit of a human calcium channel selected from the group consisting of an  $\alpha_1$ -subunit, a  $\beta$  subunit, an  $\alpha_2$ -subunit and a  $\gamma$  subunit.

14. The eukaryotic cell of claim 13, wherein the heterologous DNA encodes an  $\alpha_1$ -subunit of a human calcium 15 channel.

15. The eukaryotic cell of claim 13, wherein the  $\alpha_1$ -subunit is an  $\alpha_{1A}$  subunit, an  $\alpha_{1B}$  subunit, an  $\alpha_{1C}$  subunit or an  $\alpha_{1D}$  subunit.

16. The eukaryotic cell of claim 13, wherein the 20 heterologous DNA encodes an  $\alpha_2$ -subunit of a human calcium channel.

17. The eukaryotic cell of claim 16, wherein the calcium channel is a human skeletal muscle calcium channel or a human aortic calcium channel.

25 18. The eukaryotic cell of claim 13, wherein the heterologous DNA encodes a  $\beta$ -subunit of a human calcium channel.

19. The eukaryotic cell of claim 13, wherein the 30 heterologous DNA encodes a  $\gamma$ -subunit of a human calcium channel.

20. The eukaryotic cell of claim 13 which has a functional heterologous calcium channel that contains at least one subunit encoded by the heterologous DNA.

21. The eukaryotic cell of claim 20, wherein at least 35 one subunit encoded by the heterologous DNA is an  $\alpha_1$ -subunit of a human calcium channel.

-144-

22. The eukaryotic cell of claim 21, wherein at least two subunits are encoded by the heterologous DNA and the subunits encoded by the heterologous DNA, in addition to the  $\alpha_1$ -subunit, are a  $\beta$ -subunit or an  $\alpha_2$ -subunit.

5 23. The eukaryotic cell of claim 22, wherein the calcium channel contains at least three subunits that are encoded by the heterologous DNA.

24. The eukaryotic cell of claim 23, wherein the calcium channel contains at least four subunits encoded by  
10 the heterologous DNA.

25. The eukaryotic cell of claim 13 selected from the group consisting of HEK 293 cells, Chinese hamster ovary cells, African green monkey cells, and mouse L cells.

26. The eukaryotic cell of claim 20 selected from the  
15 group consisting of HEK 293 cells, Chinese hamster ovary cells, African green monkey cells, and mouse L cells.

27. A eukaryotic cell with a functional, heterologous calcium channel, produced by a process comprising introducing at least one RNA transcript selected from the  
20 group consisting of a first RNA which is translatable in said cell into an  $\alpha_1$ -subunit of a human calcium channel, a second RNA which is translatable in said cell into a  $\beta$ -subunit of a human calcium channel, a third RNA which is translatable in said cell into an  $\alpha_2$ -subunit of a human  
25 calcium channel, and a fourth RNA which is translatable in said cell into a  $\gamma$ -subunit of a human calcium channel.

28. The eukaryotic cell of claim 27 which is an amphibian oöcyte.

29. A method for identifying a compound that  
30 modulates the activity of a calcium channel, comprising;  
suspending a eukaryotic cell which has a functional,  
heterologous calcium channel, in a solution containing said  
compound and a calcium channel selective ion:

35 depolarizing the cell membrane of said cell; and  
detecting the current flowing into said cell,  
wherein:

-145-

the heterologous calcium channel includes at least one human calcium channel subunit encoded by DNA or RNA that is heterologous to said cell,

5 the current that is detected is different from that produced by depolarizing the same or a substantially identical cell in the presence of the same calcium channel selective ion but in the absence of said compound.

30. The method of claim 29, wherein prior to the depolarization step the cell is maintained at a holding 10 potential which substantially inactivates calcium channels which are endogenous to said cell.

31. The method of claim 30, wherein:

the cell is an amphibian oöcyte;

15 the heterologous subunits are encoded by RNA injected into the oöcyte; and

the heterologous subunits include an  $\alpha_1$ -subunit and a  $\beta$ -subunit.

32. The method of claim 31, wherein the subunits encoded by said RNA further comprise an  $\alpha_2$ -subunit, a  $\gamma$ - 20 subunit or an  $\alpha_2$ -subunit and a  $\gamma$ -subunit.

33. The method of claim 29, wherein the cell is an HEK cell and the heterologous subunit is encoded by heterologous DNA.

34. A substantially pure  $\alpha_1$ -subunit of a human calcium 25 channel encoded by the DNA of claim 1.

35. A substantially pure  $\alpha_2$ -subunit of a human calcium channel encoded by the DNA of claim 4.

37. A substantially pure  $\beta$ -subunit of a human calcium channel encoded by the DNA of claim 8.

30 38. A substantially pure  $\gamma$ -subunit of a human calcium channel encoded by the DNA of claim 12.

39. The DNA fragment of claim 1, wherein the human calcium channel is a human neural calcium channel, a human skeletal muscle calcium channel or a human aortic calcium 35 channel.

# INTERNATIONAL SEARCH REPORT

International Application No PCT/US 92/06903

## L CLASSIFICATION OF SUBJECT MATTER (If several classification symbols apply, indicate all)<sup>6</sup>

According to International Patent Classification (IPC) or to both National Classification and IPC

**IPC5: C 07 K 13/00, C 12 N 15/12, C 12 N 5/16, C 12 Q 1/00**

## II. FIELDS SEARCHED

**Minimum Documentation Searched<sup>7</sup>**

Classification System	Classification Symbols
IPC5	C 07 K; C 12 N; C 12 Q

**Documentation Searched other than Minimum Documentation  
to the Extent that such Documents are Included in Fields Searched<sup>8</sup>**

## III. DOCUMENTS CONSIDERED TO BE RELEVANT<sup>9</sup>

Category <sup>10</sup>	Citation of Document, <sup>11</sup> with indication, where appropriate, of the relevant passages <sup>12</sup>	Relevant to Claim No. <sup>13</sup>
A	WO, A1, 9113077 (THE SALK INSTITUTE BIOTECHNOLOGY/INDUSTRIAL ASSOCIATES, INC.) 5 September 1991, see the whole document ---	1-39
A	WO, A1, 8909834 (THE SALK INSTITUTE BIOTECHNOLOGY/INDUSTRIAL ASSOCIATES, INC.) 19 October 1989, see the whole document, cited in the application ---	1-39
A	WO, A1, 8907608 (NEW YORK UNIVERSITY) 24 August 1989, see the whole document ---	1-39

### \* Special categories of cited documents:<sup>10</sup>

- "A" document defining the general state of the art which is not considered to be of particular relevance
- "E" earlier document but published on or after the international filing date
- "L" document which may throw doubts on priority claim(s) or which is cited to establish the publication date of another citation or other special reason (as specified)
- "O" document referring to an oral disclosure, use, exhibition or other means
- "P" document published prior to the international filing date but later than the priority date claimed

"T" later document published after the international filing date or priority date and not in conflict with the application but cited to understand the principle or theory underlying the invention

"X" document of particular relevance, the claimed invention cannot be considered novel or cannot be considered to involve an inventive step

"Y" document of particular relevance, the claimed invention cannot be considered to involve an inventive step when the document is combined with one or more other such documents, such combination being obvious to a person skilled in the art

"Z" document member of the same patent family

## IV. CERTIFICATION

Date of the Actual Completion of the International Search

24th November 1992

Date of Mailing of this International Search Report

07 DEC 1992

International Searching Authority

Signature of Authorized Officer

EUROPEAN PATENT OFFICE

Mikael G:san Bergstrand

III. DOCUMENTS CONSIDERED TO BE RELEVANT (CONTINUED FROM THE SECOND SHEET)		
Category	Citation of Document, with indication, where appropriate, of the relevant passages	Relevant to Claim No
P,X	Proc. Natl. Acad. Sci. USA, vol. 89, January 1992, Susumu Seino et al.: "Cloning of the alpha 1 subunit of a voltage-dependent calcium channel expressed in pancreatic beta cells ", pp. 584-588, see especially the abstract, fig. 1  --	1-3,13- 15,20- 21,25- 28
P,X	Dialog Information Services, file 154, Medline, 66-92/May, accession no. 07972010, Medline accession no. 92110010, Williams M.E. et al: "Structure and functional expression of alpha 1, alpha 2, and beta subunits of a novel human neuronal calcium channel subtype", & Neuron Jan 1992, 8 (1), p71-84  --	1-39
X	Dialog Information Services, file 154, Medline 66-91/May, accession no. 07846396, Medline accession no. 91365396, Powers P.A. et al: "Assignment of the human gene for the alpha 1 subunit of the cardiac DHP-sensitive Ca <sup>2+</sup> channel (CCHL1A1) to chromosome 12p12-pter", & Genomics Jul 1991, 10 (3) p835-9  --	1-3,13- 15
X	Proc. Natl. Acad. Sci. USA, vol. 86, May 1989, A. Lee Burns et al.: "Calcium channel activity of purified human synexin and structure of the human synexin gene ", pp. 3798-3802, see especially the abstract  -----	1,4,8, 12

**ANNEX TO THE INTERNATIONAL SEARCH REPORT  
ON INTERNATIONAL PATENT APPLICATION NO.PCT/US 92/06903**

**SA 63819**

This annex lists the patent family members relating to the patent documents cited in the above-mentioned international search report.  
The members are as contained in the European Patent Office EPO file on **30/10/92**.  
The European Patent office is in no way liable for these particulars which are merely given for the purpose of information.

Patent document cited in search report	Publication date	Patent family member(s)		Publication date
WO-A1- 9113077	05/09/91	NONE		
WO-A1- 8909834	19/10/89	AU-D- EP-A- JP-T-	3548089 0424397 4503152	03/11/89 02/05/91 11/06/92
WO-A1- 8907608	24/08/89	AU-D- EP-A- JP-T- US-A-	3193289 0357730 2503202 4950739	06/09/89 14/03/90 04/10/90 21/08/90

For more details about this annex : see Official Journal of the European patent Office, No. 12/82